Evaluación Psicológica

Historia, fundamentos teórico-conceptuales y psicometría

Laura Edna Aragón Borja

2ª edición



Evaluación Psicológica

Historia, fundamentos teórico-conceptuales y psicometría

2a edición

Evaluación Psicológica

Historia, fundamentos teórico-conceptuales y psicometría

2a edición

Mtra. Laura Edna Aragón Borja
Jefa de la Unidad de Evaluación Psicológica
de la FES Iztacala, UNAM



Editorial El Manual Moderno, S.A. de C.V.
Av. Sonora 206 - 201 Col. Hipódromo, C.P. 06100 México, D.F.

Editorial El Manual Moderno, (Colombia), Ltda
Carrera 12-A No. 79-03/05 Bogotá, DC

IMPORTANTE

Los autores y la Editorial de esta obra han tenido el cuidado de comprobar que las dosis y esquemas terapéuticos sean correctos y compatibles con los estándares de aceptación general en la fecha de la publicación. Sin embargo, es difícil estar por completo seguro que toda la información proporcionada es totalmente adecuada en todas las circunstancias. Se aconseja al lector consultar cuidadosamente el material de instrucciones e información incluido en el inserto del empaque de cada agente o farmacoterapéutico antes de administrarlo. Es importante, en especial, cuando se utilizan medicamentos nuevos o de uso poco frecuente. La Editorial no se responsabiliza por cualquier alteración, pérdida o daño que pudiera ocurrir como consecuencia, directa o indirecta, por el uso y aplicación de cualquier parte del contenido de la presente obra.

Nos interesa su opinión, comuníquese con nosotros:

Editorial El Manual Moderno S.A. de C.V.

Av. Sonora 206, Col. Hipodromo, Deleg. Cuauhtémoc. 06100 México, D.F.

(52-55) 52-65-11-00

info@manualmoderno.com

quejas@manualmoderno.com

Evaluación psicológica: historia, fundamentos teórico-conceptuales y psicometría. 2a edición.

ISBN: 978-607-448-465-6 (versión impresa)
ISBN: 978-607-448-466-3 (versión electrónica)

Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. núm. 39

Para mayor información sobre

Catálogo de producto
Novedades
Distribuciones y más
www.manualmoderno.com

Aragón Borja, Laura, Edna, autor.
Evaluación psicológica : historia fundamentos teórico-conceptuales y psicometría / Laura Edna Aragón Borja. –- 2ª. edición. –- México : Editorial El Manual Moderno, 2015.

ISBN: 978-607-448-465-6 (versión impresa)
ISBN: 978-607-448-466-3 (versión electrónica)

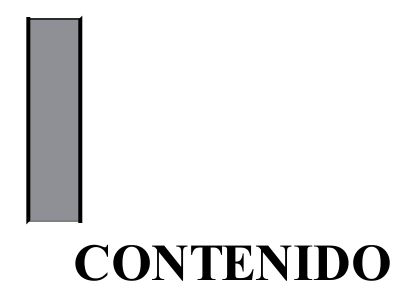
1. Pruebas psicológicas. 2. Evaluación de la conducta. 3. Psicometría. I. Título.
150.287-scdd21 Biblioteca Nacional de México

Director editorial y de producción:
Dr. José Luis Morales Saavedra

Editora asociada:
LCC Tania Uriza Gómez

Diseño de portada:
LCS Adriana Durán Arce





CONTENIDO

ACERCA DE LA AUTORA

La Maestra Laura E. Aragón Borja es Licenciada en Psicología por la Universidad Nacional Autónoma de México, cuenta con la Especialidad en Matemáticas Aplicadas a Psicología (Convenio de Beca Conacyt-ENEP Iztacala), así como con la Maestría en Modificación de Conducta, por la UNAM. Es Profesora Titular "B", de tiempo completo en la División de Investigación y Posgrado de la Facultad de Estudios Superiores Iztacala, UNAM. Tiene amplia experiencia en evaluación psicológica; desde 1994 se ha desempeñado como Jefa de la Unidad de Evaluación Psicológica, supervisando anualmente la atención de alrededor de 80 usuarios de dicha Unidad, así como la evaluación psicométrica de los aspirantes a diferentes especialidades. Actualmente imparte el Diplomado con opción a titulación "Evaluación Psicológica Integral de niños, adolescentes y adultos", en la FES Iztacala, UNAM. Este Diplomado se ha impartido consecutivamente desde 2002 a la fecha.

Es autora de una prueba para evaluar niños disléxicos y ha realizado la tipificación mexicana de la prueba española BAPAE que mide aptitudes para el aprendizaje escolar, así como de la prueba de inteligencia factor "g" escala 3 de Cattell. Así mismo, ha asesorado al Departamento de Psicología del Instituto Tecnológico de Sonora (ITSON) en la obtención de normas de puntuación del Inventario Multifásico de Personalidad Minnesota para Adolescentes (MMPI-A) en el Sur de Sonora y del Cuestionario 16 PF de Cattell también para el Estado de Sonora.

INTRODUCCIÓN

La evaluación psicológica es, además de una rama de la psicología, una disciplina científica en sí misma con objetivos básicos y de investigación, dedicada a la exploración y análisis científico de la conducta humana,cuya meta última es la toma de decisiones para la intervención.

En la actualidad, la actividad de la evaluación psicológica se ha ampliado de manera considerable: las pruebas psicológicas se usan cada vez más para evaluar a personas, grupos o instituciones, tanto para tener información más amplia acerca de sus fortalezas y debilidades como para planear un tratamiento o intervención. Las escuelas, las clínicas y las personas mismas, solicitan evaluaciones psicológicas para ayudar a aquellos sujetos que presentan problemas de aprendizaje, de conducta o emocionales, así como para seleccionar o asignar a los que no los tienen a determinadas situaciones o condiciones.

De esta manera, la evaluación psicológica cobra especial importancia, ya que se reconoce que para poder llevar a cabo una intervención psicológica exitosa y eficaz es imprescindible realizar antes una evaluación, y mientras esta evaluación se haga de la mejor manera, que aporte datos relevantes necesarios para el psicólogo, se hará una mejor planeación de la inter-vención con los subsecuentes beneficios para el paciente.

El objetivo de este libro es contribuir a la formación de psicólogos en el área de la evaluación psicológica. Para tal fin, iniciamos esta obra con la exposición de los hechos históricos más relevantes que sucedieron para que la evaluación psicológica, tal como la conocemos ahora, pudiera surgir. Para poder entender mejor nuestro presente, necesitamos conocer nuestro pasado, así, en los dos primeros capítulos se abordan los antecedentes históricos, tanto los precientíficos —aquellas evaluaciones realizadas en la antigüedad; las aportaciones de filósofos como Platón, Aristóteles y Descartes a la psicología; la doctrina del empirismo inglés— como los científicos: el surgimiento de la psicofísica, la teoría de la evolución de Darwin, los estudios de Wundt, Galton y Cattell, la elaboración de la prueba de inteligencia de Binet, la proliferación de pruebas durante la Primera y Segunda Guerras M-undiales, el establecimiento de la psicometría y algunos de los sucesos más importantes del siglo XX que contribuyeron a conformar la evaluación psicológica.

Para que una disciplina pueda ser considerada científica, necesita recoger y comparar datos que le permitan fundamentar sus teorías; el componente esencial de este proceso es la medición, uno de los conceptos que se exponen con amplitud en el capítulo 3, al igual

que otros temas fundamentales: las principales características de las pruebas psicológicas y los conceptos de testing, psicodiagnóstico y evaluación psicológica.

Una vez que la evaluación psicológica incorpora la teoría de la medida a sus procedimientos para asegurar su estatus científico, surge la psicometría —conjunto de métodos, técnicas y teorías implicadas en la medición de variables psicológicas—, dedicada al estudio de las propiedades métricas exigibles a las mediciones psicológicas y a establecer las bases para que dichas mediciones se realicen de manera adecuada. El objetivo de la psicometría es proporcionar modelos para transformar los hechos en datos, con la finalidad de asignarles valores numéricos a los sujetos sobre la base de sus respuestas. En psicología, la psicometría es la que avala que los instrumentos de medición elaborados cuenten con las garantías científicas para su uso. En este capítulo 4, explicamos la Teoría clásica de las pruebas, sobre la que descansan la mayoría de las pruebas psicológicas; las garantías científicas que toda prueba debe poseer: confiabilidad y validez, y cuáles son los diferentes tipos de puntuaciones que nos proporcionan las pruebas, así como su interpretación.

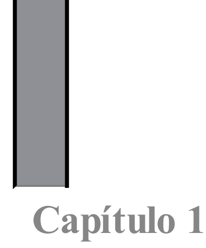
Sin embargo, antes de iniciar un proceso de evaluación con la finalidad de resguardar los derechos de las personas a las que se ofrece el servicio, debemos tener presentes las consideraciones éticas del proceso de evaluación psicológica. En el capítulo 5 se presentan los principios éticos que deben regir el comportamiento de los psicólogos: respeto a los derechos y a la dignidad de las personas, cuidado responsable, integridad en las relaciones y responsabilidad hacia la sociedad y la humanidad.

En el capítulo 6 se expone muy detalladamente la forma en que se elabora por escrito un reporte de evaluación psicológica una vez terminado el proceso de evaluación, con el objetivo de que sea de máxima utilidad tanto para el psicólogo que va a planear el tratamiento como para los usuarios del servicio para que tomen las mejores decisiones.

Finalmente en el Capítulo 7, el lector encontrará algunas sugerencias y recomendaciones para seleccionar una batería de evaluación psicológica, describiendo algunas de las pruebas más utilizadas que en la actualidad se pueden comprar en México, el objetivo de las mismas, la información que proporcionan, el tipo de puntuaciones que utilizan y para qué tipo de problemas se recomienda su aplicación.

Esperamos que esta obra cumpla sus objetivos: además de profundizar en el conocimiento de la evaluación psicológica y la psicometría, contribuir a que el psicólogo realice un ejercicio responsable de su profesión en esta importante área de la psicología.

Laura Edna Aragón Borja



Capítulo 1

ANTECEDENTES HISTÓRICOS DE LA EVALUACIÓN PSICOLÓGICA: DE LA ANTIGÜEDAD AL SIGLO XIX

En la historia de la humanidad, en todas las épocas, ha sido evidente el interés por la medición de los atributos humanos, desde los físicos —estatura, peso, color de ojos y pelo, fuerza, agudeza auditiva y visual— hasta los psicológicos —inteligencia, personalidad, motivación, entre otros. Este interés puede rastrearse desde los filósofos de la antigüedad, la filosofía del siglo XVII, en los grandes avances científicos de la medicina, la física, la química y la astronomía de los siglos XVIII y XIX, hasta la medición psicológica propiamente dicha en los albores del siglo XX con la primera prueba considerada como realmente psicométrica, elaborada por Binet, y el gran auge que la construcción de pruebas y el establecimiento de la psicometría tuvieron con la llegada de las dos guerras mundiales.

Desde el punto de vista histórico, ¿cuándo podemos afirmar que inicia la evaluación en psicología? Al respecto, Garaigordobil (1998) menciona que cuando el hombre toma conciencia de que existen variadas diferencias entre él y sus semejantes, intenta buscar procedimientos que le permitan cuantificar el grado de dichas diferencias. Así, la autora establece que para que un procedimiento se considere como evaluación psicológica debe reunir varios requisitos:

- Un evaluado y un evaluador, y la necesidad de un objetivo y un contexto para desarrollar dicha evaluación.
- Un marco teórico que permita la ubicación de las hipótesis.
- Una técnica o procedimiento de medida y la interpretación de los valores obtenidos.
- Disposición de los sistemas categoriales pertinentes que permitan la clasificación del evaluado a partir de los resultados que se hayan obtenido al aplicarle los procedimientos de medición.

El primer requisito puede parecer muy obvio, sin embargo, es necesario precisarlo porque en ocasiones se han dado diagnósticos a la distancia, sin la presencia explícita de

un evaluado y un evaluador; por ejemplo: para la fisiognomía, que establece que según los rasgos faciales de un individuo serían sus características de personalidad, bastaría sólo un retrato para evaluar a una persona y determinar que, como tiene la cara ancha en relación con la altura del rostro, tendería a presentar comportamientos agresivos. Por otro lado, en la actualidad, en un proceso de evaluación psicológica debe estar bien determinado y justificado el objetivo de evaluación: ¿qué datos necesitamos y queremos obtener al término del proceso?; no solamente aplicar un instrumento de medida porque sí, sino con antelación seleccionar los instrumentos adecuados que ayudarán a cumplir dicho objetivo, el cual debe estar inmerso en uno o más contextos específicos: escolar, familiar, clínico, laboral, etc.

El segundo requisito tiene que ver con la teoría en la que la prueba descansa y que el evaluador selecciona porque cumple sus expectativas en cuanto al tipo de datos que dicha prueba puede proporcionar. Si por ejemplo, lo que se pretende es confirmar o descartar un diagnóstico de ansiedad, se debe seleccionar una prueba que mida ansiedad y que desde el marco teórico en que dicha prueba está construida nos proporcione datos para confirmar las hipótesis planteadas.

Las pruebas psicológicas son instrumentos de medición con reglas muy claras en cuanto a la obtención y transformación de las puntuaciones que proporcionan y que nos van a permitir interpretar los valores encontrados a la luz de la teoría sobre la que la prueba descansa. De igual forma, estos mismos valores están dispuestos en diferentes categorías para poder comparar al evaluado con el grupo de referencia y así determinar que se encuentra en la norma, o bien, por arriba o por debajo de ella, y las implicaciones que esto tiene para la toma de decisiones.

De esta manera, al considerar los requisitos anteriores, se puede hablar de antecedentes precientíficos y antecedentes científicos de la evaluación psicológica (Garaigordobil, 1998). Dentro de los primeros estarían las evaluaciones realizadas en la antigüedad y las derivadas y nutridas por la gran influencia en psicología de la filosofía (Platón, Aristóteles, Descartes; empirismo inglés) y la medicina (Gall). En los antecedentes científicos, encontramos a la psicofísica, la teoría de la evolución y los estudios de Wundt, Galton y Cattell.

Antecedentes precientíficos

Los historiadores señalan que las primeras evaluaciones de aptitudes datan del año 2200 a. C., en China, donde los funcionarios de gobierno eran evaluados a petición del emperador para determinar su competencia profesional (Chafee, 1985; citado en Gregory, 2001). Estas pruebas fueron aplicadas y modificadas a lo largo de los siglos y se convirtieron en exámenes escritos durante la Dinastía Han (202 a. C. a 200 d. C.), para adquirir su forma final hacia el año 1370. Esta evaluación consistía en pedir a los evaluados, como primera prueba, que pasaran un día y una noche en una pequeña cabina aislada, donde tenían que realizar composiciones escritas sobre temas asignados; si aprobaban esta prueba, continuaban con otro tipo de evaluaciones extenuantes y

rigurosas durante tres días y tres noches. Finalmente, a los que aprobaban se les aplicaba la última etapa de evaluación en Pekín, donde quienes triunfaban obtenían el puesto de mandarín. De acuerdo con Gregory (2001) y Martínez (1996), este programa de exámenes se amplió poco a poco, y hacia 1370, para ocupar puestos en el gobierno, incluía la evaluación de las aptitudes en equitación, música, aritmética, escritura (donde la belleza de la caligrafía tenía un peso importante, ya que era esencial para una buena comunicación); conocimiento de las leyes civiles, los ritos o ceremonias y los principios de Confucio. Este tipo de exámenes fue finalmente eliminado en 1905 y reemplazado por requisitos educativos, tales como poseer un título universitario.

El filósofo Platón (427 – 347 a. C.) menciona, específicamente en *La República*, que no hay dos personas iguales, que cada una difiere de las otras en sus dotes naturales y que, por tanto, una persona puede ser apropiada para desarrollar una profesión y otra para desenvolverse en alguna totalmente distinta; él consideraba que si bien todos los seres humanos básicamente son iguales, sin embargo diferirían en sus cualidades y aptitudes; en cada persona estarían desarrolladas por naturaleza unas capacidades más que otras, por lo tanto él proponía que de acuerdo a dichas capacidades se estableciera una República formada por tres clases de ciudadanos, esto es, una división social del trabajo: los filósofos, que serían los encomendados a gobernar por poseer la virtud de la sabiduría; el pueblo, formado por los artesanos y los labradores, trabajarían en actividades productivas; los guerreros, fuertes y valientes, con la virtud de la fortaleza, estarían destinados a velar por el orden y la defensa. Así, proponía llevar a cabo una serie de funciones para evaluar las aptitudes militares con el propósito de seleccionar a los mejores soldados. Anastasi (1980) considera que estas acciones propuestas por Platón con la finalidad de evaluar los rasgos que se consideraban esenciales para determinar la destreza militar representan la primera prueba de aptitud de la cual existe memoria descrita sistemáticamente, cuyo objetivo era la asignación de los individuos a tareas específicas para los que estaban dotados.

Aristóteles (384 – 322 a. C.), otro gran filósofo griego, discípulo de Platón, se percató de la existencia de las variaciones individuales. En sus escritos, plantea que hay diferencias en las características mentales y morales de los individuos de acuerdo con el nivel social, la raza y el sexo, y las atribuye a factores innatos. En su *Ética nicomáquea*, Aristóteles plantea que la plena excelencia sólo podría ser alcanzada por el varón adulto y maduro perteneciente a la clase alta, y la cual nunca sería alcanzada por las mujeres, los niños, los bárbaros o no griegos, y los asalariados o trabajadores manuales; de esta manera, él consideraba naturales las relaciones de poder del hombre sobre la mujer, los padres sobre los hijos y los amos sobre los esclavos.

Consideraba que aunque un hombre se propusiera ser el mejor, no lo conseguiría, a menos que estuviera en su naturaleza, y que al comparar a unos hombres con otros, se evidenciaría que en la gimnasia, la medicina, la construcción, la navegación, las ciencias en general y cualquier actividad no científica existirían diferentes niveles de aptitud. Asimismo, consideraba que los hombres poseían determinados rasgos o cualidades —como la ira, la audacia, la inmoralidad— de manera diferencial, ya sea en exceso, en

término medio o de modo deficiente (Anastasi, 1980).

De igual manera, Aristóteles proponía el análisis del carácter y los hábitos psicológicos de las personas con base en sus características corporales (fisiognomía). Se atribuye a Aristóteles un tratado llamado *"Physiognomonica"*, en el cual se relacionan los signos corporales con el carácter y las facultades mentales de las personas. Estas teorías fisiognómicas ya habían sido anteriormente postuladas tanto por Empédocles (450 a. C.), como por Hipócrates (400 a. C.), en sus conocidas Doctrinas Humorales de Personalidad. Estas Doctrinas establecen que el cuerpo humano está compuesto por cuatro humores: sangre, bilis amarilla, bilis negra y flema, que a su vez corresponden a los siguientes elementos de la naturaleza, respectivamente: aire, tierra, fuego y agua, y que de acuerdo al humor que predomine en cada individuo, éste presentará alguno de los cuatro tipos de temperamento y personalidad: sanguíneo y optimista, melancólico y triste, colérico e iracundo, y flemático y apático. Esta postura influyó posteriormente en los trabajos tipológicos de Kretschner, en 1925, y su discípulo Sheldon, en 1949, sobre el temperamento (Pervin, 1984): tipo pícnico o endomorfia (tendencia a la obesidad, que caracteriza a los individuos como amables, dependientes, joviales y extrovertidos, de fácil expresión y emoción; muestra amor por el confort, tiende a evitar el dolor y es dependiente de la aprobación social), tipo atlético o mesomorfia (gran desarrollo muscular y esquelético, que distingue a los hombres como agresivos, con gran energía física, coraje, poder, aserción, cierta indiferencia y tolerancia al dolor físico, necesidad de acción y ejercicio) y tipo asténico o ectomorfia (escaso desarrollo de las estructuras musculares, frágil estructura ósea, que caracteriza a los individuos como introvertidos, tímidos, románticos, solitarios, pensativos y muy reflexivos).

Por su lado, René Descartes (1596 – 1650) en su *Discurso del método*, establece una separación radical entre el mundo material (los cuerpos), regidos por una necesidad mecánica de moverse, y el alma, que es una propiedad exclusivamente humana caracterizada por el pensamiento inteligente; los cuerpos se mueven y obran de acuerdo a leyes mecánicas, las almas pueden decidir por libre albedrío. El cuerpo humano es una máquina y está unida al alma a través de la glándula pineal, por lo que alma y cuerpo se encuentran en interacción constante, considerando que las impresiones sensoriales, a través de los nervios y los espíritus animales en ellos contenidos, hacen sentir a las almas de los sujetos distintas pasiones o emociones que ella se encargará, a través de su albedrío, de fortalecer o rechazar. De esta forma, Descartes plantea que no es posible estudiar al hombre de manera objetiva, por lo que hay que utilizar la intuición para acceder a los datos de la conciencia. En una carta dirigida a Hyperaspistes, en 1641, él plantea que no se puede establecer una verdad universal sobre el comportamiento humano al modo geométrico, es decir, como la lograda en el conocimiento matemático, por lo que aquí se deduce el carácter subjetivo de la conducta humana (Rocha, 2004). Su filosofía influyó posteriormente en los enfoques fenomenológicos y humanistas, que consideran que no es posible llevar a cabo evaluaciones objetivas cuantificadas, y subyace en la creación de las pruebas subjetivas o proyectivas.

En Inglaterra, entre los siglos XVII y XVIII, se desarrolla la doctrina filosófica del

empirismo, la cual considera a la experiencia como la única fuente válida de conocimiento. Para John Locke (1632 - 1704), al igual que para Descartes, el objeto del conocimiento humano son las ideas, pero a diferencia de él, considera que no hay más ideas en nuestro entendimiento que las que provienen de la experiencia y que el único conocimiento que los humanos pueden poseer es un conocimiento a posteriori. Locke estuvo fuertemente influido por el filósofo también inglés Francis Bacon (1561 – 1626), quien decía que la única forma de llegar a dominar la naturaleza sólo se conseguía observándola, que era preciso estudiar lo que sucede, anotar los resultados y hacer experimentos para reforzar las observaciones realizadas y que sólo sobre esa base se podrían obtener interpretaciones acertadas científicamente (Savater, 2009). Para los empiristas, la única fuente de conocimiento verdadero del universo era la experiencia sensible, lo que podía verse, oírse, sentirse, palparse u olerse, y lo que se podía inferir de las relaciones entre estas experiencias sensibles. El empirismo inglés es el antecedente de la doctrina filosófica del positivismo, iniciada en la primera mitad del siglo XIX, que identifica al verdadero conocimiento como aquel derivado del método científico o conocimiento positivo, es decir, a los hechos cuya verdad fuera incuestionable, cuya evidencia se apoyara en los sentidos y cuya validez estuviera garantizada por el método científico; a su vez el positivismo es el antecedente de la corriente psicológica del conductismo, la cual considera solamente como objeto de estudio a la conducta del sujeto externa, objetiva y observable.

La influencia médica en esta primera etapa precientífica se debe al médico Franz Joseph Gall (1758 – 1828), quien planteó que las facultades psíquicas estaban localizadas en el cerebro y que podrían ponerse al descubierto al analizar las protuberancias craneales. Desde que era un niño, Gall había estado observando las distintas formas del cráneo de las personas, y estaba firmemente convencido de que tanto las diferentes formas como el tamaño deberían estar relacionados con las características personales; él era localizacionista, es decir, pensaba que las funciones mentales se encontraban ubicadas en zonas específicas del cerebro, y que dado que el cráneo es el recipiente del cerebro, observando sus formas y protuberancias, se podrían determinar las características de personalidad de los sujetos. Bajo estos supuestos se desarrollará la frenología —la cual mediante "mapas" localiza en el cráneo distintas características psicológicas, como la emotividad, la inteligencia, la agresividad, etc.—, antecedente de la evaluación neuropsicológica.

Antecedentes científicos: siglos XVIII y XIX

Ecuación personal

En 1796, un ayudante de astrónomo llamado Kinneybrook discrepaba siempre de Nevil Maskelyne —el astrónomo real del Observatorio Astronómico de Greenwich, en Inglaterra— en las mediciones que realizaba. Estas mediciones consistían en anotar el tiempo que tardaban en pasar ciertas estrellas en una marca realizada con un hilo en el

telescopio. Kinneybrook fue despedido porque sus mediciones diferían en promedio más de medio segundo después de las realizadas por Maskelyne, lo cual provocaba errores considerables en los cálculos astronómicos (Nunnally, 1973; Anastasi, 1980).

En 1816, el astrónomo alemán Friedrich Bessel se interesó por el incidente de Kinneybrook y decidió medir las diferencias en segundos de las estimaciones entre dos observadores. Bessel publicó datos sobre parejas de observadores entrenados y descubrió divergencias en cualquier par de observadores, así como diferencias de medición de cada sujeto comparado consigo mismo en distintos momentos, Bessel llegó así a la conclusión de que los errores en los cálculos de los astrónomos se debían a las personas, es decir, a los tiempos de reacción de los astrónomos a la hora de observar por el telescopio, lo cual, por lo tanto, daba como resultado la diferencia en los cálculos.

Este trabajo es considerado como el primer registro publicado sobre las diferencias individuales, lo que fue conocido posteriormente como la ecuación personal, es decir, la tendencia característica a estimar las observaciones asignándoles una determinada cantidad, ya sea por exceso o por defecto (Anastasi, 1980; Nunnally, 1973).

Este incidente tuvo tres consecuencias afortunadas: primero, reconocer que las personas difieren en sus juicios y que tales diferencias individuales pueden ser medidas, lo cual reafirmaba que en los atributos humanos hay diferencias individuales; segundo, que el ser humano como un instrumento de medición puede cometer errores, y tercero, que como cualquier otra herramienta de medición, los instrumentos de la ciencia psicológica (incluido el ser humano) también poseen un margen de error que hay que considerar y tener en cuenta, así como incorporar dichos errores al análisis de los datos, lo que a la larga sería el fundamento del concepto de confiabilidad.

Curva normal: Quetelet

Durante el siglo XIX, ya los astrónomos habían estado utilizando la curva normal de Gauss para describir el comportamiento de los errores de medición, pero fue Adolphe Quetelet (1796 – 1874), un astrónomo, matemático, naturalista y sociólogo belga, quien pensó que esta distribución normal podía también ser usada para explicar algunos aspectos de la conducta humana; él afirmaba que era posible identificar regularidades de comportamiento que caracterizaran al hombre promedio y así contrastarlas contra el comportamiento de individuos aislados a través de las desviaciones de los mismos sobre una curva normal (Nunnally, 1973). Quetelet se dedicó a reunir una gran cantidad de información en poblaciones europeas: datos sobre la locura, los suicidios, la delincuencia, el número de niños en las familias, el número de niños nacidos en distintos años, así como características físicas de las personas.En una ocasión tomó las medidas de la estatura de cien mil soldados franceses, en otra, las medidas torácicas de poco más de 5,700 soldados escoceses; encontró que al graficar estas medidas, seguían una distribución de curva normal. Llegó a la conclusión de que una distribución normal de las características humanas mostraba el error de la naturaleza de los seres humanos, es decir, que la meta de la naturaleza era crear al hombre normal o promedio, y que los accidentes

o errores de la naturaleza se podrían poner de manifiesto en los extremos de la curva normal.

Psicofísica: estudios de Weber y Fechner

Históricamente, se podría considerar que el primer problema de medición psicológica abordado de manera sistemática fue el estudio de las sensaciones. Uno de los primeros intentos que se hizo para emplear el método experimental en psicología se dirigió hacia la solución del problema de la relación entre las sensaciones percibidas por las personas al cambiar los estímulos que se les presentaban. Para Miller (1983), la posibilidad de medir las sensaciones fue la prueba de que la psicología podría algún día convertirse en ciencia. Si las sensaciones eran factibles de ser medidas, podría entonces también medirse cualquier fenómeno de la percepción, analizándolo mediante su reducción a los componentes elementales y mensurables de la experiencia.

Poco después del interés por el estudio de la ecuación personal, tanto los filósofos como los primeros psicólogos se interesaron por estudiar el umbral de conciencia, a lo que se le llamó limen, que es el punto o el umbral en el cual se llega a adquirir conciencia visual, auditiva o cinestésica de la existencia de un objeto. Por ejemplo, si un sonido o una luz son demasiado débiles o tenues para poder escucharse o verse, decimos que se encuentra por debajo del umbral o limen de la percepción auditiva o visual; a estos estímulos que no pueden detectarse se les llama subliminales, y a los que sí se pueden detectar supraliminales. Al punto en el que el ojo, el oído o el sentido del tacto perciben un objeto se le llama limen absoluto. Como dato interesante, podemos agregar que esta manera de medir los umbrales diferenciales todavía se utiliza; por ejemplo, en los exámenes de agudeza auditiva, para determinar los umbrales auditivos se le presentan al sujeto sonidos en diferentes niveles de intensidad y se le pide que mencione cuándo empieza a oír el sonido, o bien, cuándo deja de escucharlo.

Para Garret (1975), la psicofísica comenzó con la obra del fisiólogo alemán Ernst Heinrich Weber (1795 – 1878), quien como profesor de anatomía en la Universidad de Leipzig, de 1829 a 1834, publicó una serie de experimentos sobre las sensaciones cutáneas y cinestésicas. Weber estaba interesado en descubrir con cuánta exactitud las personas podrían apreciar pequeñas diferencias de peso cuando dos pesos se levantaban con la mano, o bien, cuando se dejaban libremente en la superficie de la piel, así como cuál sería la diferencia más pequeña entre dos líneas que pudiera ser discernida por el ojo humano; a esto se le llamó limen diferencial, que comprende tanto la diferencia mínima perceptible como la diferencia mínima no perceptible.

La primera, la diferencia mínima perceptible, es la diferencia que el ser humano puede percibir entre dos estímulos —visuales (por ejemplo, una luz con cierta intensidad), auditivos (un sonido con determinados decibeles) o cinestésicos (un objeto con determinado peso)— que comienzan siendo iguales, para luego uno de ellos variar de intensidad y dejar al otro fijo, hasta que la persona pueda determinar que uno es más o menos intenso que el otro. Por otro lado, en la diferencia mínima no perceptible, se le

presentan al sujeto dos estímulos muy distintos y poco a poco se modifica uno de ellos hasta que el individuo reporta que ya no ve, oye o siente la diferencia.

Weber encontró que la diferencia mínima perceptible entre dos objetos tiende a estar en relación constante con la intensidad inicial del estímulo (Nunnally, 1973). Por ejemplo, un individuo podría determinar que hay una diferencia entre un peso de 30 onzas y otro de 31 o 29 onzas, pero no sentir la diferencia entre un peso de 50 onzas y otro de 51; para este último caso, la diferencia de peso debería ser de 1.65 onzas, esto es, el objeto debería pesar 51.65 o bien de 48.35 onzas para poder sentir la diferencia con respecto a uno de 50. Al realizar diferentes experimentos, Weber concluyó que lo que percibíamos no era la diferencia real entre dos objetos, sino la razón que esta diferencia guarda con las magnitudes de los dos objetos comparados. Así, estableció la llamada Ley de Weber, en donde la diferencia mínima perceptible (dmp) es igual a la intensidad del estímulo (S) por una constante (c), esto es: dmp = cS. En el caso anterior, la constante sería 0.033 (30 X 0.033 = 0.99; 30+0.99 = 30.99; 30-0.99 = 29.01; 50X0.033 = 1.65; 50+1.65 = 51.65; 50-1.65 = 48.35). Matemáticamente, la Ley de Weber se expresa como dR/R = C, donde R es el estímulo y dR, el incremento por el que R debe aumentarse con objeto de producir un cambio apenas perceptible, esto es, se establece la relación que existe entre los juicios sensoriales y las intensidades de los estímulos (Garret, 1975). En conclusión, la Ley de Weber indica que la cantidad de excitación que se necesita para distinguir entre una sensación y otra, está en una relación constante con la sensación inicial; si la cantidad de excitación se aumenta poco a poco, la primera sensación no cambia al principio, y para que el sujeto pueda percibir el aumento, esto es, experimentar una sensación diferente, se requiere un aumento importante proporcional a la excitación primera. Como bien considera Garaigordobil (1998), con los estudios de Weber sobre la medición de los umbrales diferenciales en distintos niveles de estimulación se abría la posibilidad de medir lo subjetivo.

Gustav Theodor Fechner (1801 – 1887) continuó el estudio de la Ley de Weber, de hecho, la mayor parte de la psicofísica de Fechner, puede considerarse como la culminación de los resultados de Weber (Kantor, 2005). Fechner ingresó a la Universidad de Leipzig en 1818 y se quedó en ella como profesor e investigador hasta su muerte; aunque se graduó en medicina, su desarrollo profesional fue principalmente como físico y filósofo (Nunnally, 1973; Garret, 1975). Sus especulaciones filosóficas se ocupaban de la presencia de la conciencia en todas las cosas, incluso en plantas y objetos inanimados. Como punto de partida de sus estudios de psicofísica, retomó los experimentos de Weber sobre juicios humanos y en particular sobre la diferencia mínima perceptible. En la Ley de Weber descubrió un medio para estudiar cuantitativamente la relación entre los mundos físico y mental. Sus investigaciones partían del postulado de que si bien la sensación no puede medirse directamente, era legítimo preguntarle a un individuo si la sensación existe o no, o bien, si una sensación era más intensa que otra, es decir, consideraba que como las respuestas encubiertas a los estímulos no son directamente accesibles para la investigación psicológica, pero las respuestas manifiestas sí lo son, habría que llegar a las sensaciones por medio del reporte verbal del sujeto, lo cual se

puede considerar como el antecedente remoto del método del cuestionario.

En su obra *Elementos de Psicofísica*, publicada en 1860, Fechner consideraba a la psicofísica como el estudio de las relaciones que hay entre los sucesos del mundo físico y las respuestas psicológicas, en donde el mundo físico estaba representado por los estímulos físicos y el mundo psíquico por las sensaciones que esos estímulos provocan en el organismo. Definía a la psicofísica como la ciencia de la relación funcional entre el cuerpo y la mente, entre lo material y lo espiritual, entre lo físico y lo psíquico (Nunnally, 1973; Garret, 1975).

Al método creado por Fechner —de presentar de manera precisa y sistemática los estímulos físicos para suscitar las respuestas manifiestas— se le llamó psicofísico y es una de sus principales contribuciones a la medición de juicios humanos en psicología. Su método de comparación por pares se utiliza actualmente en los exámenes de agudeza visual, donde se le pide al sujeto que mencione si existe diferencia entre una dioptría u otra que el oftalmólogo le presenta (¿con cuál ve mejor, con ésta o con ésta?). Este método también se utiliza en la comparación y clasificación de perfumes, licores, tabacos y, en psicología, en algunas pruebas de orientación vocacional, en donde el sujeto tiene que elegir entre tres estímulos aquello que más le gusta y lo que menos le gusta. Otro método psicofísico creado por Fechner y que se refiere a la graduación de los estímulos en diferentes intensidades es el antecedente remoto de las escalas de actitudes tipo Likert, utilizadas hoy en día para evaluar los atributos psicológicos.

La obra de Fechner, a quien se le considera el padre de la psicofísica, es importante porque demostró de manera concluyente que los problemas de la psicología pueden analizarse mediante métodos cuantitativos, y además se ve reflejada en el amplio desarrollo posterior de las pruebas psicológicas. En palabras de Kantor (2005), si bien Fechner es el fundador de un culto que decía medir la relación entre el cuerpo y la mente, desde el punto de vista de la psicología científica se le reconoce como auténtico estudioso de la conducta, y por tanto, como líder del avance científico.

Establecimiento de la psicología como ciencia experimental: Wundt

Derivado de los estudios de Weber sobre los umbrales diferenciales y de Fechner con la creación de sus métodos psicofísicos, se abre el camino hacia la medición de lo subjetivo mediante técnicas de autoinforme. Si bien estos avances contribuyeron al progreso de la medición en psicología, el antecedente científico fundamental del psicodiagnóstico es, según Garaigordobil (1998), la constitución de la psicología como disciplina experimental en 1879, cuando Wilhem Wundt establece en Leipzig, Alemania, el primer laboratorio de psicología experimental con la finalidad de estudiar fenómenos psicológicos con la metodología propia de la investigación fisiológica de esa época; su objetivo era determinar la relación de los fenómenos psíquicos con su sustrato orgánico, particularmente cerebral. Lo que Wundt entendió por psicología experimental, según Kantor (2005), fue la psicología fisiológica, esto es, una psicología que utilizaba las técnicas de la fisiología, estudiando actividades tales como la sensación, la percepción, la

atención y el tiempo de reacción. Wundt desempeñó un papel decisivo en la constitución de la psicología experimental, anexándole la fisiología y la anatomía; su meta fue elaborar una psicología admitiendo solamente hechos y recurriendo a la experimentación y a la medición (Mueller, 2004). En dicho laboratorio se formaron psicólogos de la talla de Brentano, Ebbinghaus, Dilthey, Hull, James y el propio Wundt; casi la mitad de los directores de los primeros laboratorios de psicología de Estados Unidos se formaron en el laboratorio de Wundt (Hogan, 2004).

Para Wundt, la única forma de estudiar la conciencia era reduciéndola a sus elementos más simples, entender la forma en que estos elementos se hallan conectados para poder determinar las leyes de esa conexión; al igual que los químicos habían analizado la materia reduciéndola a átomos y los anatomistas reduciendo los sistemas vitales a células, así entonces se debía analizar la mente reduciéndola a las sensaciones y sentimientos elementales que la constituyen (Miller, 1983). Así, las aportaciones de Wundt a la evaluación psicológica son (Garaigordobil, 1998; Hogan, 2004):

1) Concentra la experimentación en psicología en la estandarización de las condiciones y la precisión de la medición, y enfatiza la importancia de realizar las observaciones de todos los sujetos bajo condiciones tipificadas.
2) Contextualiza a la psicología dentro de las ciencias naturales como ciencia empírica y, por tanto, separada de la filosofía.
3) Postula el método experimental para el estudio de la conciencia, ya que todo elemento de ésta se deriva directa o indirectamente de la sensación y toda idea compleja es una combinación asociativa de elementos más simples.
4) En aras de la objetividad, limita sus observaciones a los procesos más elementales de la conciencia, tales como tiempos de reacción, sensaciones auditivas y visuales, percepción, psicofísica, reacciones motrices simples y asociación, al considerar que los fenómenos mentales más complejos como el pensamiento no podrían ser estudiados en el laboratorio.

Teoría de la evolución de Charles Darwin

El biólogo inglés Charles Darwin (1809 – 1882) tuvo una gran influencia no sólo en el campo de la biología, sino también en la psicología. En 1859, publica su magna obra *El origen de las especies*, en la que concebía a los animales y las plantas en un estado de cambio constante en el que progresan gradualmente de una forma a otra para adaptarse a los medios cambiantes. También establece que las diferentes especies, mediante un mecanismo de selección natural, se adaptan y subsisten, y otras se extinguen, a lo que llamó "la supervivencia del más apto" (Nunnally, 1973).

La teoría de la evolución de Darwin favoreció el estudio de las diferencias individuales en psicología, ya que si estas diferencias en las plantas y animales explicaban su diversa capacidad para adaptarse y sobrevivir, las mismas tendrían en los seres humanos una importancia funcional; y si las plantas y los animales heredaban características

ancestrales, algunas de las características individuales de los seres humanos podrían ser explicadas sobre la misma base de la herencia (Nunnally, 1973).

Mientras Wundt buscaba los principios universales y generales que gobiernan la mente, los psicólogos interesados en la evolución comenzaron a catalogar todas las formas en que las mentes podían diferir entre sí (Miller, 1983). El papel de Darwin y su teoría de la evolución resultó de gran relevancia, ya que las diferencias individuales no solamente empezaron a concebirse como algo totalmente normal —dejando de considerarse como errores de medición o errores de la naturaleza, como lo postulaba Quetelet—, sino incluso a aceptarse como algo adaptativo (Garigordobil, 1998).

Primeros estudios sobre las diferencias individuales: Galton

El objetivo principal de los psicólogos formados experimentalmente era encontrar las leyes de la conducta humana, es decir, formular descripciones generalizadas de la conducta. En este sentido, cuando en sus experimentos encontraban diferencias individuales, ignoraban los resultados, o bien, los aceptaban como un mal necesario que limitaba la aplicación de las generalizaciones. Así, cuando se observaba que un individuo reaccionaba de modo distinto a otro en condiciones idénticas, se consideraba como una forma de error, ya que impedía establecer las deseadas generalizaciones. Esta actitud de considerar las diferencias individuales como errores prevaleció entre los psicólogos formados en el laboratorio de Wundt, actitud fundamentada en los hallazgos de Quetelet.

Las investigaciones sobre psicofísica de Weber y Fechner y el enfoque evolucionista de Darwin influyeron mucho en los trabajos acerca del estudio de las diferencias individuales de Francis Galton (1822 – 1911), en Inglaterra, y posteriormente de McKeen Cattell (1861 - 1944), en Estados Unidos.

Galton, primo segundo de Charles Darwin, era un caballero inglés, rico e independiente, pero con gran interés en la ciencia (un intelectual autodidacta). Sin tener formación universitaria, emprendió el estudio de una amplia gama de temas que abarcaban la criminología (fotografió delincuentes para tratar de caracterizar los rasgos criminales), la biología, la antropología (para comprender las religiones primitivas y el significado del ídolo para éstas, creó una religión propia) y desde luego la psicología (para estudiar la paranoia, imaginó que todos hablaban de él, llegó incluso a creer que hasta los caballos tramaban algo en su contra). Inventó algunos dispositivos científicos para medir los sentidos como un silbato para medir las frecuencias más altas que puede percibir el oído humano: el pito supersónico (silbato con el tono más agudo que puede percibirse, con el cual llamaba a los perros y éstos se apartaban de sus incrédulos amos); la barra de Galton para la discriminación visual de la longitud, la fotografía compuesta (superposición de una fotografía sobre otras), entre otros (Nunnally, 1973).

Sin embargo, sus estudios no se limitaron a estos fenómenos sensoriales. El interés principal de Galton era comprobar la heredabilidad de las características individuales y la medición de los rasgos humanos. Postulaba que no sólo las características físicas prominentes se heredan, sino que también las características personales —las aptitudes y

los rasgos de personalidad, la inteligencia o el "genio", las tendencias delictivas y los trastornos psicológicos— se transmitían de padres a hijos por herencia directa.

En 1869, con su obra *Heredity Genius*, intentó demostrar mediante historias familiares la herencia de talentos específicos en diferentes áreas: música, literatura, etc. Por ejemplo, en un estudio descubrió que en 30 familias donde ambos padres tenían talento artístico, el 64% de sus hijos también lo presentaban, contra el 21% de los hijos de 150 familias en las que ninguno de los padres tenía talento. Después de este libro, publicó otros con el mismo tema: *English Men of Science*, en 1874, y *Natural Inheritance*, en 1889 (Nunnally, 1973; Anastasi, 1980; Garaigordobil, 1998; Hogan, 2004).

Derivado de la hipótesis de que las características personales se heredaban, Galton intentó crear el estudio de la eugenesia, cuyo propósito explícito sería el mejoramiento de la raza humana y la creación de un mundo de seres humanos superiores por medio del control de la reproducción, en donde se alentaría a las personas con talentos especiales —gran inteligencia, genialidad en música, literatura, liderazgo militar o político— a que se casaran y tuvieran hijos y se desalentaría a los individuos menos dotados a que los tuvieran.

Sin embargo, Galton comprendió que antes de emprender un programa de eugenesia de tal magnitud, era necesario tener un mayor conocimiento de aquellas características individuales que realmente se heredan y, más aún, antes de proceder a estudiar la transmisión de estas características mediante la herencia, se necesitaba medirlas de manera objetiva y en grandes números; ésta fue la segunda gran contribución de Galton a la psicología: el intento de medir las características psicológicas (Nunnally, 1973; Garret, 1975; Anastasi, 1980; Hogan, 2004). Y como él mismo menciona: "…hasta que un fenómeno de cualquier rama del conocimiento no haya sido sometido a la medida y asignación de números, no podemos asumir que tenga el status y la dignidad de ciencia…" (Galton, 1879, p. 147; cit. en Martínez, 1996).

Galton estaba firmemente convencido de que la capacidad mental de cualquier tipo dependía de los procesos psicológicos más elementales. Partía de la hipótesis de que las personas que, al igual que los animales, tuvieran los sentidos más agudos serían las mejor dotadas y por tanto las de mejores capacidades intelectuales. Publica así en 1883 su obra *Inquiries into Human Faculty and its Development* —punto de partida de las pruebas mentales—, donde enfatiza las diferencias individuales de acuerdo con las facultades mentales; considera que: "La única información que llega hasta nosotros en relación a los acontecimientos exteriores pasa, al parecer, por la avenida de nuestros sentidos; y cuanto más perciban la diferencia nuestros sentidos, mayor será el campo sobre el cual pueden actuar nuestro juicio y nuestra inteligencia." (Galton, 1883, p. 27; citado en Anastasi, 1977, 1980). En la misma obra, mencionaba que los retrasados mentales tienden a ser deficientes en la capacidad para discriminar el calor, el frío y el dolor, por lo que esto reforzaba aún más su convicción de que la capacidad discriminativa sensorial debería ser superior en aquellos mejor dotados intelectualmente. Mediante la medida de los procesos sensoriales, Galton esperaba llegar a un cálculo del nivel intelectual de los sujetos. Así, bajo esta premisa, emprendió el primer programa de administración de mediciones de

características humanas en su Laboratorio Antropométrico durante la Exposición Internacional sobre Salud, que se efectuó en Londres en 1884, donde utilizaba pruebas de agudeza sensorial, discriminación, tiempos de reacción y mediciones fisiológicas simples, como ventanas hacia la mente. En dicho laboratorio, por una cantidad de tres peniques, le tomaba a cada visitante medidas de una gran variedad de pruebas físicas y sensoriales: dimensiones de la cabeza, dimensión de los brazos extendidos, altura, peso, capacidad respiratoria, fuerza muscular en ambas manos, rapidez de golpeteo, agudeza auditiva y determinación de umbrales de audición, agudeza visual, discriminación visual de longitudes, tiempo de reacción auditiva y visual, series graduadas de pesos para medir la discriminación cinestésica, sentido del color, precisión al dividir una línea en dos y tres partes iguales, error al estimar la apertura de un ángulo de 90 grados y otro de 60, etc. Cuando la exposición fue clausurada, el laboratorio se trasladó al Museo South Kensington de Londres, en donde funcionó durante seis años más, en el curso de los cuales fueron evaluadas de 10 000 a 17 000 personas, de los que se conservan todavía en la actualidad 7 500 registros (Nunnally, 1973; Anastasi, 1977; Martínez, 1996; Garaigordobil, 1998; Gregory, 2001; Muñiz, 2003; Hogan, 2004).

Para analizar los datos obtenidos, recurrió a los métodos estadísticos y determinó promedios y medidas de varianza; al tratar de relacionar las diferentes variables estudiadas, creó el diagrama de distribución bivariada e indujo a Karl Pearson (1857 – 1936) a inventar el coeficiente de correlación; es el primer investigador en psicología que utilizó y desarrolló métodos estadísticos para analizar los datos provenientes de sus mediciones (Nunnally, 2003; Muñiz, 2003; Hogan, 2004). Otro de sus discípulos, Charles Spearman (1863 – 1945), desarrolló la teoría de la confiabilidad de las pruebas, el análisis factorial y el modelo estadístico de puntaciones conocido como Teoría Clásica de las Pruebas o, como se le conoce generalmente, Teoría Clásica de los Tests (TCT) (Martínez, 1996). Por sus grandes contribuciones al campo de la medición en psicología y la evaluación psicológica, a saber: la postulación acerca de la heredabilidad de las características psicológicas, la demostración de que las diferencias individuales no sólo existen, sino que se pueden medir de manera objetiva, el comienzo del movimiento de las pruebas mentales y el uso de procedimientos estadísticos en psicología, se le puede considerar como el padre de la psicología diferencial.

Primeras pruebas mentales: McKeen Cattell

James McKeen Cattell (1861 – 1944), psicólogo estadounidense, fue también un personaje clave en la configuración de la evaluación psicológica. Realizó sus estudios bajo la dirección de Wundt, en Leipzig, por lo que tuvo una formación experimental rigurosa. En 1886, para su doctorado concluyó una disertación sobre las diferencias individuales en el tiempo de reacción; su tesis se tituló *The Time Taken up by Cerebral Operations* (Garret, 1975). Dos años más tarde, en Cambridge, fue discípulo de Galton, quien reforzó su interés por el estudio de las diferencias individuales (Anastasi, 1980; Hogan, 2004). Rechaza abiertamente la introspección como forma de abordar los

fenómenos psicológicos y sostiene que el objetivo prioritario de la disciplina es el uso de procedimientos objetivos en el estudio de las ejecuciones de los sujetos, sin atribuir a éstas inferencias o atributos subyacentes (Garaigordobil, 1998).

A su regreso a los Estados Unidos, en la Universidad de Columbia, Cattell se concentra en la creación de laboratorios de psicología experimental y en la elaboración y aplicación de pruebas. Al igual que Galton, Cattell consideraba que era posible obtener una medida del funcionamiento intelectual mediante la medición de los procesos elementales, tales como las funciones sensoriales, perceptivas y motoras. Creó una batería de 50 pruebas cuyo objetivo, de acuerdo con los planteamientos de Galton, era la determinación del rango, exactitud y naturaleza de las facultades psicológicas más elementales, así como la posibilidad de reunir suficiente material para determinar los factores que regulan el desarrollo de estas facultades, sus conexiones y sus perturbaciones, por lo que se convirtió en el primer psicólogo que planteó el uso de baterías de pruebas para la evaluación psicológica.

En un artículo que escribió en 1890 para la revista *Mind*, "Mental Tests and Measurements", acuñó el término de "test mental", el cual se emplearía para caracterizar el área durante los siguientes 50 años, y describió una serie de 10 pruebas que aplicaba anualmente a estudiantes universitarios con el objetivo de determinar su nivel intelectual y pronosticar su desempeño académico. Las pruebas eran (Anastasi, 1980; Garaigordobil, 1998; Gregory, 2001, Hogan, 2004):

1. Fuerza de la presión ejercida con la mano, medida con un dinamómetro.
2. Tasa o velocidad de movimiento de la mano en el trayecto de una distancia de 50 cm.
3. Umbral de tacto de dos puntos (diferencia mínima no perceptible donde dos puntos todavía se perciben como separados).
4. Sensibilidad al dolor: grado de presión necesaria para sentir dolor con una punta de goma contra la frente.
5. Percibir la diferencia de peso de cajas en apariencia idénticas, desde los 100 hasta los 110 gr.
6. Tiempo de reacción para el sonido: rapidez con que puede moverse una mano al iniciarse una señal auditiva.
7. Tiempo para nombrar colores: tiempo que lleva enumerar los colores de una franja de 10 papeles de color.
8. Dividir en dos partes iguales una línea de 50 cm.
9. Estimación de un tiempo de 10 segundos.
10. Memoria o número de letras repetidas después de escucharlas una sola vez.

Cattell tuvo varios discípulos que después se convirtieron en psicólogos famosos (Thorndike y Woodworth); uno de ellos, Clarke Wissler, obtuvo las puntuaciones de las pruebas mentales y de las calificaciones escolares de 300 estudiantes de la Universidad de Columbia y del Colegio Barnard, con el objetivo de demostrar que los resultados de

las pruebas mentales podían pronosticar el desempeño académico. Los resultados fueron expuestos por Wissler, en 1901, en un artículo que publicó en *Psychological Review*, "The Correlation of Mental and Physical Tests", en el cual mostró que la correlación entre la ejecución en las pruebas mentales y las calificaciones escolares era muy baja, e incluso en algunas pruebas era negativa (la más alta fue de 0.16 en la prueba de memoria) (Gregory, 2001).

Las pruebas de Cattell fueron un fracaso colosal como indicadores de la aptitud intelectual. Los análisis de los datos mostraron la nula correlación entre la ejecución de los alumnos en dichas pruebas y su nivel intelectual; la ejecución de los estudiantes evaluados presentaba poca consistencia de una prueba a otra y poca o ninguna relación con valores independientes del nivel intelectual, basados en las estimaciones de los profesores o en las calificaciones académicas (Anastasi, 1980; Muñiz, 2003; Hogan, 2004).

Con la publicación de los fallidos resultados encontrados por Wissler, los psicólogos experimentales abandonaron el uso de las pruebas sensoriales y motoras como indicadores de la capacidad intelectual, lo cual abonó el camino para la elaboración de pruebas psicológicas que midieran procesos mentales superiores. Para Gregory (2001), la reacción de los psicólogos al iniciar el siglo XX fue concluir a regañadientes que Galton había estado equivocado en sus intentos por inferir capacidades complejas a partir de capacidades simples. Sin embargo, el abandono del desarrollo de pruebas que midieran la capacidad intelectual no duraría mucho tiempo; en 1905, en Francia, Alfred Binet elaboraría la primera prueba propiamente psicométrica de la historia.

Referencias

Anastasi, A. (1977). *Tests psicológicos*. España: Aguilar.

Anastasi, A. (1980). *Psicología diferencial*. España: Aguilar.

Fernández-Ballesteros, R. (1996). *Introducción a la evaluación psicológica.* Salamanca: Pirámide.

Garaigordobil, L. M. (1998). *Evaluación psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro*. España: Amarú Ediciones.

Garret, H. E. (1975). *Las grandes realizaciones de la psicología experimental*. México: FCE.

Gregory, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México: El Manual Moderno.

Hogan, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.

Kantor, J. R. (2005). *La evolución científica de la psicología.* México: Trillas

Martínez A. R. (1996). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.

Miller, G. A. (1983). *Introducción a la Psicología*. Madrid: Alianza Editorial.

Mueller, F. L. (2004). *Historia de la psicología.* México: Fondo de Cultura Económica.

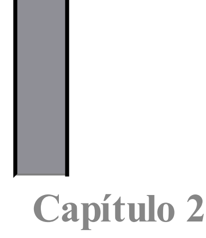
Muñiz, J. (2003). *Teoría clásica de los tests*. Madrid: Pirámide.

Nunnally, J. C. (1973). *Introducción a la medición psicológica*. Buenos Aires: Paidós.

Pervin, I. A. (1984). *Personality: Theory and Research*. New York: John Willey and Sons, Inc.

Savater, Fernando (2009). *Historia de la Filosofía. Sin temor ni temblor.* España: Espasa

Rocha, H. L. (2004). "La idea de hombre en la filosofía cartesiana (Una proyección hacia la individualidad)". *Revista Digital Universitaria.* 10 de abril de 2004. http://www.revista.unam.mx/vol.5/num3/art17/art17.html [Consulta: 11 de abril de 2004].



Capítulo 2

ANTECEDENTES HISTÓRICOS DE LA EVALUACIÓN PSICOLÓGICA EN LA ÉPOCA MODERNA: SIGLO XX

Como vimos en el capítulo anterior, el estudio de las diferencias individuales se extendió de Inglaterra, con los trabajos de Galton, a los Estados Unidos, con la aplicación de baterías de pruebas mentales por Cattell, principalmente de tipo sensorial y motor. Sin embargo, con la publicación de Wissler en 1901 (Gregory, 2001), se llegó a la conclusión de que estas pruebas no medían la inteligencia, lo cual echaba por tierra la hipótesis de que la capacidad sensorial estaba estrechamente relacionada con la inteligencia.

Al comenzar el siglo XX, Francia era el centro del humanismo, del interés por la medicina y el bienestar social; expresiones como “libertad”, “igualdad” y “fraternidad”, reflejaban el interés de ese país por ayudar al oprimido, al enfermo y al deficiente mental (Nunnally, 1973). En ese contexto, en 1904, el Ministerio de Educación Pública de París le encomendó a un equipo de médicos, educadores e investigadores (entre los que estaban Alfred Binet y Theodor Simon) que decidiesen qué medidas educativas deberían emprenderse para distinguir al niño perezoso o inadaptado de aquél que no tenía capacidad de aprender o con deficiencia mental, esto es, para detectar a aquellos niños que no podían beneficiarse de la educación regular, con la finalidad de proporcionarles una pedagogía apropiada. Esta comisión concluyó que para tal fin deberían utilizarse exámenes médicos y educativos. Sin embargo, como era evidente que se necesitaba además un medio para poder asignar a los niños a una u otra categoría, se le solicitó a Binet, que en ese tiempo era director del Laboratorio de Psicología Fisiológica de la Sorbona, y a su colaborador Simon que desarrollaran una herramienta práctica dirigida únicamente a ese propósito, lo que dio lugar a la primera escala formal para evaluar la inteligencia en niños: la Binet-Simon Scale (Nunnally, 1973; Garret, 1975; Gregory, 2001; Hogan, 2004). En este segundo capítulo sobre la historia de la evaluación psicológica, abordaremos la elaboración de la prueba de inteligencia de Binet, la proliferación de pruebas durante la Primera y Segunda Guerras Mundiales, el establecimiento de la psicometría como rama de la psicología científica y algunos sucesos importantes recientes en la configuración de la evaluación psicológica.

La primera prueba psicométrica de inteligencia: Binet

Binet (1857 -1911) nació en Niza, Francia. Aunque su formación original fue en Derecho —después realizó estudios superiores en medicina y ciencias naturales—, sus intereses lo dirigieron hacia la psicología; sus primeros trabajos en esta disciplina versaron sobre el uso del hipnotismo en el tratamiento de los trastornos mentales (Hogan, 2004). En 1895, Binet y su ayudante en la Sorbona, Victor Henri, publicaron un artículo sobre la revisión de los trabajos alemanes y estadounidenses acerca de las diferencias individuales: "La Psychologie individuelle"; en él criticaban este tipo de trabajos por considerar que utilizaban en demasía pruebas de tipo sensorial y que se concentraban indebidamente en capacidades sencillas y especializadas; planteaban que la medición de la inteligencia debería abordarse mediante procesos psicológicos o mentales superiores: "…no son las sensaciones, son las facultades psíquicas superiores lo que hace falta estudiar, son ellas las que juegan el puesto más importante y la psicología individual deberá prestar mucha más atención a ellas" (Binet y Henri, 1895, p.416; citado en Garaigordobil, 1998). Para Binet, la capacidad intelectual —esos procesos mentales superiores— se debería revelar en el razonamiento, la imaginación y el juicio, los que, junto con la adaptabilidad, serían los contribuyentes más importantes de la inteligencia. Garret (1975) menciona que Binet consideraba que el intelecto se componía de juicio, sentido común, iniciativa, capacidad de adaptación, conciencia de las propias capacidades, adaptación inmediata de la propia conducta a una meta o fin definidos, y persistencia en la ejecución de una tarea emprendida.

Binet es el primero que plantea un concepto claro de diagnóstico psicológico como sinónimo de evaluación de la inteligencia, al distinguirlo del examen médico y de la evaluación pedagógica (análisis de los aprendizajes). Destaca que una prueba psicológica debería contar con tres requisitos: tareas simples, aplicación rápida y resultados contrastables por otros observadores (Garaigordobil, 1998). Con dichos requisitos, sobre todo el último que hace referencia a la posibilidad de que los resultados puedan ser repetidos y contrastados por diferentes evaluadores, Binet está sentando las bases para la estandarización de las pruebas psicológicas objetivas.

Así, en colaboración con Simon, Binet elaboró su primera prueba de inteligencia en 1905, la cual ya no se ocupaba específicamente de evaluar funciones sensoriales o motoras como las de Galton y Cattell, en lugar de eso, proponía la utilización de pruebas mentales en las que habría que determinar la capacidad del niño para comprender y razonar acerca de los objetos de su ambiente, estudiar la memoria, la naturaleza de las imágenes mentales, la imaginación, la atención, la comprensión, los sentimientos estéticos y morales, y aún la fuerza de voluntad, por lo que su prueba, aunque todavía incluía reactivos sensoriales y perceptivos, abarcaba una proporción mucho mayor de contenido verbal, con reactivos tales como nombrar objetos, completar oraciones o comprender preguntas (Nunnally, 1973; Anastasi, 1977; Garaigordobil, 1998). Binet aprobaba la inclusión de los diferentes reactivos basándose en la observación y la práctica, más que

en la teoría; observaba a los niños y dejaba que su ejecución decidiera qué reactivos eran adecuados y cuáles inaplicables en cada grupo de edad. Esta escala de 1905 ordena los reactivos seleccionados por Binet en orden de dificultad creciente, y así cada niño realizaba todos los ítems que podía hasta que le resultaran demasiado difíciles; posteriormente se comparaba el nivel al que llegaba el niño con el nivel que alcanzaban los niños de su misma edad. Esta primera escala constaba de 30 reactivos que iban de lo más simple a lo más complejo y fueron aplicados a 50 niños de entre tres y 11 años, así como a algunos con retraso mental. Si bien las pruebas o reactivos no se agrupaban de acuerdo con la edad, Binet indicó qué pruebas debería esperarse que pasara un niño normal de una edad determinada. Los reactivos eran los siguientes (Sattler, 1996; Jenkins y Paterson, 1961, cit. en Gregory, 2001):

1. Seguir con los ojos un objeto en movimiento (coordinación visual)
2. Tomar un objeto pequeño cuando se le pone al sujeto en contacto con la palma o el dorso de su mano
3. Tomar un objeto pequeño cuando se le pone a su vista y se le alienta verbalmente a que lo alcance
4. Reconocer la diferencia entre un cuadro de chocolate y uno de madera
5. Búsqueda de comida cuando se interpone una pequeña dificultad: ingerir un cuadro de chocolate envuelto en papel
6. Obedecer órdenes sencillas e imitar ademanes simples
7. Darle al experimentador objetos familiares nombrados como taza, llave, cuerda, así como tocarse partes de su cuerpo: cabeza, nariz, orejas
8. Señalar objetos representados en láminas
9. Nombrar objetos representados en dibujos. Estas tres últimas pruebas eran el límite superior para los niños normales de tres años de edad.
10. Comparar dos líneas de longitud notablemente desigual
11. Repetir tres dígitos inmediatamente después de su presentación verbal
12. Comparar dos pesos (3 y 12 gramos, 6 y 15, y 3 y 15)
13. Mostrar susceptibilidad a la sugestión. Se le pide un objeto que no se encuentra entre los presentes, se le pide que diga dónde está un... (se usa una palabra sin sentido) y se le dan a comparar dos líneas de igual longitud.
14. Definir palabras comunes (casa, caballo, tenedor, mamá)
15. Repetir oraciones de 15 palabras inmediatamente después de escucharlas
16. Explicar la diferencia entre dos objetos comunes (papel y cartón, mosca y mariposa, madera y vidrio). Esta prueba separa de manera efectiva a los niños de 5 y 7 años; límite para los 7 años.
17. De un total de 13 objetos presentados simultáneamente, mencionar de memoria el mayor número posible en 30 segundos
18. Reproducir de memoria dos diseños mostrados durante 10 segundos
19. Repetir de memoria series mayores a tres dígitos
20. Decir en qué se parecen objetos familiares: amapola (roja) y sangre; hormiga,

mosca, mariposa y pulga; periódico, etiqueta y fotografía

21. Comparar dos líneas de longitud desigual pero muy cercana, por ejemplo, una de 30 cm con 15 líneas que van desde los 31 a los 35 cm
22. Ordenar cinco cubos de igual tamaño por peso (15, 12, 9, 6 Y 3 gramos)
23. Indicar cuál de los cinco cubos anteriores ha sido retirado. Esta prueba es el límite más probable para aquellos niños con deficiencia mental
24. Encontrar palabras que rimen con una palabra dada, después de que se han presentado ejemplos
25. Encontrar la palabra faltante en una oración sencilla
26. Dados tres sustantivos, elaborar una oración
27. Responder 25 preguntas sobre cuestiones de la vida cotidiana (comprensión). Por ejemplo, ¿qué es lo que haces cuando tienes sueño?, ¿por qué es mejor continuar lo que se está haciendo que comenzar algo nuevo?
28. Decir la hora que sería si se invirtieran las manecillas pequeña y grande del reloj, a las 2:46 y a las 6:20
29. Dibujar cómo se vería una hoja de papel doblada en cuartos, a la que se le ha quitado una de las piezas dobladas
30. Explicar diferencias entre términos abstractos como querer y respetar, estar triste y estar aburrido

Esta primera escala fue seguida de una edición aumentada y corregida en 1908 y por una tercera y última revisión en 1911, publicada un poco antes de la muerte de Binet. La primera escala de 1905 presentaba a los niños una serie de preguntas que iban desde las más simples, que cualquier niño podría contestar correctamente, hasta las más complejas, que indicaban un nivel de capacidad intelectual superior. Al aplicar esta prueba a los niños de tres a 11 años de edad, Binet comprobó que ciertos reactivos eran más difíciles de lo previsto para algunos niños de determinada edad y otros, demasiado fáciles, así como algunos que no discriminaban adecuadamente las diferencias de capacidad. De esta manera, en 1908, se realizó una revisión de los reactivos; esta escala contenía 58 reactivos y en ella las pruebas se asignaron a cada grupo de edad de la siguiente manera: si entre el 50% y el 75% de los niños de un determinado grupo de edad pasaban una prueba, Binet determinaba que era apropiada para ese grupo de edad; si todos o casi todos fracasaban en ella, la consideraba evidentemente demasiado difícil para esa edad y la colocaba en una edad superior; en tanto que si todos los niños pasaban la prueba con éxito, quedaba claramente establecido que era demasiado fácil para esa edad y la colocaba en una edad inferior. El nivel más alto en que un niño podía contestar correctamente se denominaba *nivel mental*, término que posteriormente fue cambiado por **edad mental**, de esta manera se comparaba la edad cronológica de cada niño evaluado con su edad mental para determinar su nivel de atraso. En esta segunda escala de 1908, las pruebas se ordenaron por primera vez por grupos de edad, conteniendo cada edad de cuatro a ocho pruebas. Garret (1975) considera que este método para determinar en qué edades colocar las diferentes pruebas o reactivos, aunque un tanto burdo, poseyó

la virtud de fundarse en pruebas experimentales y no en impresiones subjetivas. En relación con las técnicas psicométricas actuales, este método facilitó que los resultados de la aplicación de su prueba, esto es, los puntajes obtenidos, siguieran una distribución normal, ya que, si recordamos las propiedades de la curva normal, entre la media y una desviación estándar se encuentra casi el 70% de los casos medidos, y el resto, distribuido por igual en los extremos de cada lado de la curva.

La tercera escala, la de 1911, representa tres años más de trabajo en la escala de 1908, y fue el modo en que Binet dejó finalmente su prueba; tiene 54 reactivos dispuestos por grupos de edad de la siguiente manera (Garret, 1975, pp. 279-280):

3 años:

1. Señala la nariz, los ojos y la boca
2. Repite dos cifras
3. Enumera objetos en un cuadro
4. Pronuncia su apellido
5. Repite una oración de seis sílabas

4 años:

1. Nombra su sexo
2. Nombra la llave, el cuchillo y el centavo
3. Repite tres cifras
4. Compara dos líneas

5 años:

1. Compara dos pesos
2. Copia un cuadrado
3. Repite una oración de 10 sílabas
4. Cuenta cuatro monedas
5. Une las mitades de un rectángulo dividido

6 años:

1. Distingue entre la mañana y la tarde
2. Define palabras familiares en función del uso
3. Copia un rombo
4. Cuenta 13 centavos
5. Distingue dibujos con caras bonitas y feas

7 años:

1. Muestra la mano derecha y la oreja izquierda
2. Describe un cuadro
3. Ejecuta tres órdenes dadas simultáneamente
4. Cuenta el valor de seis monedas, tres de las cuales son dobles
5. Nombra cuatro colores principales

8 años:

1. Compara dos objetos de memoria
2. Cuenta de 20 a cero
3. Nota omisiones en cuadros
4. Da el día y la fecha
5. Repite cinco cifras

9 años:

1. Da cambio de 20 centavos
2. Define palabras conocidas en términos superiores al uso
3. Reconoce todas las monedas (nueve)
4. Nombra los meses en orden
5. Responde o comprende preguntas fáciles

10 años:

1. Ordena cinco cubos en orden de peso
2. Copia dos dibujos de memoria
3. Critica declaraciones absurdas
4. Contesta o comprende preguntas difíciles
5. Usa tres palabras dadas en no más de dos oraciones

12 años:

1. Resiste la sugestión en cuanto al largo de líneas
2. Compone una oración con tres palabras dadas
3. Nombra 60 palabras en tres minutos
4. Define tres términos abstractos
5. Descubre el sentido de una oración desordenada

15 años:

1. Repite siete cifras
2. Encuentra tres rimas para una palabra determinada en un minuto
3. Repite una oración de 26 sílabas

4. Interpreta cuadros
5. Interpreta hechos

Adulto:

1. Resuelve la prueba de cortar papel
2. Reordena un triángulo en la imaginación
3. Da diferencia entre parejas de términos abstractos
4. Explica tres diferencias entre un presidente y un rey
5. Explica la idea principal de un fragmento que se le ha leído

Tanto en esta última revisión como en sus dos escalas anteriores, Binet incluyó pruebas relativamente breves, no muy difíciles de administrar, que abarcaban una amplia gama de actividades mentales y que no dependían directamente de una información escolar específica, atributos que él consideraba debería poseer una buena prueba de inteligencia. Si nos detenemos a analizar los reactivos, podemos observar las diferentes habilidades y capacidades que miden, entre otras: atención, seguimiento de instrucciones, información tanto escolarizada como de conocimientos de la vida cotidiana, comprensión verbal, amplitud de vocabulario, memoria, resolución de problemas de aritmética básica, reconocimiento de objetos, formación de conceptos, coordinación visomotora, pensamiento abstracto, juicio estético y sentido común. Fue tanta la popularidad de esta prueba, que en menos de 10 años después de esta última revisión, la escala Binet-Simon se usaba en EUA, Canadá, Inglaterra, Australia, Nueva Zelanda, Sudáfrica, Bélgica, Alemania, Suiza, Italia, Rusia y China, y había sido traducida al japonés y al turco (Garret, 1975). Podemos considerar que esta prueba elaborada por Binet es el comienzo de la evaluación de la inteligencia como la conceptualizamos en la actualidad, es todavía el prototipo de las pruebas modernas que miden inteligencia y, aún más, es la primera prueba realmente psicométrica en la historia de la psicología.

La prueba de Binet-Simon despertó un interés creciente en países como Inglaterra, pero sobre todo en Estados Unidos, donde se hicieron varias traducciones y revisiones. Una de las más importantes fue la que realizó Terman, en 1916, en la Universidad de Stanford, a la que se le conoció como el Test de Stanford-Binet, la cual permitió la selección de reactivos apropiados para cada edad y la determinación de las normas para la interpretación de los puntajes. Con el fin de obtener la uniformidad en los procedimientos, se empleó medio año para entrenar examinadores y otro medio año en la supervisión de las pruebas administradas, se ensayaron nuevos reactivos (90 en total, 36 más que la escala de Binet de 1911), se colocaron en grupos más apropiados de edad y se mejoraron las instrucciones. El principal grupo experimental estuvo integrado por mil niños en edad escolar y clase social media (Nunnally, 1973; Garret, 1975).

Ya en 1912, William Stern había planteado que la diferencia absoluta entre la edad o nivel que un niño alcanzaba al ser evaluado con la escala Standford-Binet, y su edad cronológica, para determinar el grado de atraso o adelanto presentado, no era del todo

apropiada, ya que esta diferencia no significa lo mismo en diferentes edades (por ejemplo no es lo mismo una diferencia de dos años, a la edad de seis, que a la edad de 15), por lo que él sugería que se dividiera la edad mental del niño entre la edad cronológica, para tener una mejor medida del funcionamiento relativo de cada persona en comparación con el grupo de personas de su misma edad, a lo que se llamó Coeficiente Intelectual, y que Terman, en 1916, sugirió a su vez multiplicar por 100 para eliminar las fracciones; fue también el primero en utilizar la abreviatura IQ (en español CI) para denotarlo (Nunnally, 1973; Garret, 1975, Gregory, 2001). Terman y Merrill presentaron en 1937 otra importante revisión del Test de Stanford-Binet, aplicado a más de 3 000 niños, de dos a 15 años de edad de todas las regiones de Estados Unidos, para obtener normas y así establecer una medida estandarizada de la inteligencia (Nunnally, 1973; Garret, 1975). En 1960 se le hizo otra revisión, en 1972 se publicaron nuevas normas y en 1986 Thorndike, Hagen y Sattler publicaron una cuarta edición de la Escala Standford Binet en un formato de escala por puntos.

Podemos resumir que las aportaciones de Binet a la evaluación psicológica son: el interés por la evaluación de las funciones psicológicas superiores, la consideración del examen psicológico como independiente del médico y del pedagógico y, desde luego, su aportación más grande, la construcción de la primera prueba realmente psicométrica de inteligencia. Como bien lo menciona Hogan (2004), la demarcación entre las raíces de las pruebas psicológicas y su florecimiento —de Francia en el continente europeo a los Estados Unidos en América—, se da con las escalas de Binet. Esta prueba se convirtió en la definición de referencia de la inteligencia humana y quizás en el símbolo más distintivo de la aportación de la psicología al mundo moderno.

Las dos guerras mundiales: primeras pruebas colectivas de inteligencia y de aspectos emocionales

En 1917, cuando Estados Unidos incursiona en la Primera Guerra Mundial (1914 – 1918), se considera la necesidad de evaluar a una gran cantidad de individuos, 1 750 000 reclutas, con la finalidad de detectar su capacidad intelectual y admitirlos en el ejército, seleccionar a aquellos de más alta capacidad para asignarlos a puestos de responsabilidad y descartar a los de baja capacidad que, más que ayuda, pudieran representar un problema. De esta manera, se integró una comisión especial de la Asociación Norteamericana de Psicología (*American Psychological Association*), encabezada por R. M. Yerkes y donde también colaboraba Terman, para elaborar una prueba colectiva de inteligencia, ya que evaluar a grandes masas de personas con la prueba de Binet hubiera resultado una empresa imposible. Arthur Otis, estudiante de doctorado de Terman, ya había trabajado en 1918 en una prueba colectiva de inteligencia para su disertación doctoral, la cual fue adoptada tras algunas revisiones realizadas por dicha comisión. Esta prueba, a la que se le llamó *Army Alfa Test* (hoy se le conoce como Test Alfa), era breve y contenía ocho subpruebas con elementos verbales: 1) seguimiento de instrucciones orales, 2) razonamiento aritmético, 3) juicio en materias prácticas, 4) pares

de sinónimos-antónimos, 5) oraciones en desorden, 6) completar series numéricas, 7) analogías e 8) información general. Posteriormente, se construyó una versión no verbal de la misma, el Test Beta, para aplicarse a personas analfabetas o bien con poco dominio del idioma inglés (inmigrantes), el cual requería resolver laberintos, rompecabezas geométricos, completar series utilizando los símbolos X-O, ponerle ciertos símbolos a cifras, completar cuadros o matrices, verificar números y analizar ilustraciones. El contenido y formato de ambas pruebas fueron la base del desarrollo posterior tanto de las pruebas colectivas como de las pruebas no verbales (Nunnally, 1973; Gregory, 2001; Hogan, 2004). Para Garret (1975), el gran número de soldados evaluados constituyó el primer gran experimento psicológico de "ingeniería" humana. Como resultado de la aplicación de las pruebas Alfa y Beta, se exentaron a 8 000 hombres del servicio militar por su capacidad limitada, 10 000 fueron asignados a batallones de trabajo y otros servicios que requerían un bajo nivel de inteligencia y 10 000 fueron enviados a batallones de desarrollo especial para su observación y entrenamiento; por otra parte, casi la tercera parte de los examinados era analfabeta o leía tan mal que fue considerada como tal.

Asimismo, durante la Primera Guerra Mundial, los psicólogos, además de pruebas colectivas de inteligencia, intentaron también satisfacer la necesidad de las pruebas colectivas de neuroticismo e inestabilidad emocional —ya que había demasiados reclutas para poder entrevistarlos individualmente—, con el objetivo de descartar para el servicio militar a aquellos individuos con problemas en estas áreas. Así, en 1914, R. S. Woodworth construyó un cuestionario de 116 reactivos, a los cuales se respondía "sí" o "no": *The Personal Data Sheet*, que esencialmente formulaba la clase de preguntas que se utilizarían en una entrevista para medir ajuste personal (adaptación del individuo al hogar, a la escuela y con sus amigos), por ejemplo: *¿Las ideas pasan por tu mente de manera que no te permiten dormir? ¿Se te consideraba como un niño malo? ¿Te sientes perturbado por un sentimiento de que las cosas no son reales? ¿Tienes un fuerte deseo de suicidarte?* También durante esta guerra, el psiquiatra suizo Hermann Rorschach elaboró, en 1921, el Test Rorschach o de las Manchas de Tinta para la medición de la personalidad, en el que le pedía al sujeto que describiera e interpretara 10 manchas de tinta. La prueba tiene su origen en la psicología freudiana y servía para conocer los rasgos e impulsos más profundos que determinan la conducta manifiesta del individuo (Nunnally, 1973; Fernández-Ballesteros, 1996; Gregory, 2001; Hogan, 2004).

Nuevamente otra gran guerra —la Segunda Guerra Mundial (1939 – 1945)— habría de crear la necesidad de contar con métodos para clasificar a los individuos; sin embargo, esta vez la necesidad se presentó en una escala mayor, tanto en cantidad como en diversidad: se debían ubicar a más de 10 000 000 de personas en una amplia gama de ocupaciones nuevas en los campos de la navegación aérea (p. ej. pilotos, ingenieros de vuelo, copilotos, bombarderos), de la electrónica y de la meteorología, y de ocupaciones tales como operadores de radio, oficinistas, conductores, cocineros y otras tareas especializadas (Garret, 1975). Ahora, en lugar de pedirle a los psicólogos que construyeran pruebas colectivas de inteligencia o personalidad, se les requería baterías de

pruebas y pruebas de aptitudes específicas: aptitud mecánica, destreza motora y manual, velocidad y coordinación visual, entre otras, con el fin de escoger personal que desempeñara tareas especiales o labores de entrenamiento.

Una vez terminada la guerra, la crisis social exigió acciones de reconstrucción nacional y de readaptación personal, económica, social y laboral. Uno de los problemas que se presentaron fue la integración de los excombatientes a la vida civil; sólo en Estados Unidos, se reporta que alrededor de 1 500 psicólogos estuvieron prestando servicios como clínicos en labores de evaluación y de intervención terapéutica. Otra de las consecuencias de la guerra fue la emigración masiva de científicos europeos hacia los Estados Unidos, con lo cual este país comenzó a tener un progreso y poder a nivel científico sobre Europa y Latinoamérica, predominio que incluyó también a la psicología en general y la evaluación psicológica en particular (Garaigordobil, 1998).

A decir de Nunnally (1973), la Segunda Guerra Mundial tuvo como consecuencia la instrumentación del mayor programa de pruebas jamás emprendido, lo cual permitió la ampliación de los fundamentos lógicos y los métodos de medición en psicología y el avance de la psicología de las pruebas. Este esfuerzo contrasta con el de las décadas anteriores; ahora no sólo se pedía la elaboración de nuevas pruebas, sino cumplir con los requerimientos metodológicos y científicos para el desarrollo y uso de los métodos de medición en psicología y para la construcción de sus instrumentos, lo que llevaría a la fundación de la psicometría como rama de la psicología científica y como requisito indispensable para la elaboración de pruebas.

Como observamos en la historia de la evaluación psicológica en general y de la elaboración de pruebas en particular, han sido las necesidades sociales imperantes de cada época las que han determinado la construcción de pruebas específicas para detectar, en un primer paso, la existencia de un problema, en segundo, determinar la intensidad o gravedad del mismo y en tercero, con los datos obtenidos de la aplicación de pruebas, orientar hacia su solución. Así, en la época actual, nos hemos visto en la necesidad de crear pruebas para medir dislexia, trastorno por déficit de atención e hiperactividad, trastornos alimentarios, ansiedad y depresión en la infancia, entre otros.

Establecimiento de la psicometría como rama de la psicología: Thurstone

Como ya revisamos, a partir de los cuestionamientos de Binet sobre las pruebas utilizadas por Galton y Cattell para medir las capacidades intelectuales, de la elaboración de su prueba de inteligencia y, sobre todo, a raíz de las dos guerras mundiales, se constituye el mayor programa de elaboración y de evaluación de los atributos psicológicos por medio de pruebas; sin embargo, los cuestionamientos ahora se enfocaban en el uso y abuso de las pruebas y en la metodología utilizada para su elaboración. Si la psicología era ya considerada como una disciplina científica, sus instrumentos, al igual que los de toda ciencia, deberían cumplir con los estándares científicos para considerarse como tales, por tanto, estar enmarcados dentro de una teoría de la medida, esto es, su elaboración tenía que estar supeditada a un modelo

matemático que explicase los resultados, además de contar con las garantías indispensables de confiabilidad y validez.

Así, en congruencia con esta nueva etapa en la elaboración de pruebas —la correspondiente a la fundamentación metodológica—, en 1931, aparece el libro de L. L. Thurstone: *The Reability and Validity of Tests* y, en 1935, se establece la Sociedad Psicométrica Americana presidida por él mismo, con el objetivo de estimular el desarrollo de la psicología como ciencia cuantitativa y racional, y la cual destaca que la psicometría debería estar al servicio de los problemas psicológicos mediante el estudio de las propiedades métricas exigibles a las mediciones psicológicas. De esta asociación surge también la revista Psychometrika (Muñiz, 2003).

En 1946, S. S. Stevens propone en un artículo en la revista Science ("On the Theory of Scales of Meassurement") diferentes escalas de medición (nominal, ordinal, intervalo y razón), así como sus propiedades, relaciones e implicaciones en el uso de técnicas estadísticas. En 1947, Thurston publica otro libro: *Análisis Factorial Múltiple*, el cual será de gran utilidad para el desarrollo y construcción de pruebas. Finalmente, en 1954, se publican las primeras recomendaciones técnicas para el uso de las pruebas: *Technical Recommendations for Psychological Tests and Diagnostic Techniques*.

Como podemos observar, la historia de la evaluación psicológica por medio de pruebas psicométricas es relativamente reciente, apenas poco más de un siglo, y el establecimiento de una metodología que garantizara las propiedades lo es aún más (apenas tres cuartos de siglo). No ahondaremos más sobre la psicometría, ya que le dedicamos un capítulo completo más adelante.

A continuación presentamos una cronología de los hechos más relevantes en la historia de la evaluación psicológica, desde la creación del término "test mental" hasta 1994 (Fernández-Ballesteros, 1996; Sattler, 1996; Garaigordobil, 1998; Gregory, 2001; Muñiz, 2003), y específicamente en México, algunos hechos desde 1896, en que se imparte por primera vez una cátedra de psicología, otras fechas importantes como 1950 y 1973 y del periodo que va de 1980 a 2011.

Cronología de hechos relevantes en la historia de la evaluación psicológica y la psicometría

1890: James McKeen Cattell acuña el término "test mental".
1896, México: Se imparte por primera vez una asignatura de psicología en México en 1896, en la Escuela Nacional Preparatoria.
1901: Clark Wissler descubre que los instrumentos utilizados por Galton y Cattell no se correlacionan con las calificaciones universitarias.
1904: Karl Pearson formula la Teoría de la Correlación.
1905: Binet publica la primera prueba psicométrica de inteligencia.
1909: Henri H. Goddard traduce la Escala Binet-Simon del francés al inglés.
1914: Stern introduce el término Coeficiente Intelectual (CI).
1915: Aparece la prueba de laberintos de Porteus, la cual pretende la evaluación de las

capacidades mentales libre de influencias culturales.
1916: Primera adaptación americana de la prueba de Binet por Terman (Escala Stanford-Binet).
1918: Evaluación de inteligencia colectiva de los soldados americanos por medio de las pruebas *Army Alfa y Army Beta.*
1920: Woodworth construye la primera prueba colectiva de ajuste personal, el *Personal Data Sheet*, para la evaluación colectiva de soldados americanos.
1921: Rorschach publica su *Psychodiagnostik*, en donde presenta su Prueba de las Manchas de Tinta.
Cattell, Thorndike y Woodworth fundan la primera editorial importante de pruebas: la *Psychological Corporation.*
1926: Goodenough edita la Prueba del Dibujo de la Figura Humana.
1927: Strong publica el Inventario de Intereses Vocacionales.
1931: Thurstone publica su libro *The Reliability and Validity of Tests.*
1933: En Berlín, se prohíben oficialmente las pruebas psicológicas.
1935: Vygotsky formula el concepto de "zona de desarrollo próximo", antecedente de la evaluación de estrategias cognitivas.
Se funda la Sociedad Psicométrica Americana.
Buros publica su primera revisión de las pruebas (*Mental Measurements Yearbook*).
1936: En la Unión Soviética, el Comité Central del Partido Comunista prohíbe las pruebas mentales como instrumentos de medida. Fueron objeto de ataques debido a que se consideraba en ese país, de acuerdo a una filosofía marxista, que la meta de las pruebas era seleccionar individuos.
Se publica la primera escala de medida del desarrollo social, la Vineland Social Maturity Scale de Doll.
1938: Murray y Morgan presentan su teoría de personalidad y su prueba proyectiva, el Test de Apercepción Temática (TAT).
Lauretta Bender publica la Prueba Gestáltico Visomotor de Bender.
Arnold Gessell presenta su Escala de Maduración.
Thurstone publica la Prueba de Habilidades Mentales Primarias.
1939: Wechsler presenta la primera edición de su prueba de inteligencia para adultos (WAIS).
1943: Hathaway y McKinley publican el primer inventario clínico de evaluación de entidades nosológicas psiquiátricas, el *Minnesota Multiphasic Personality Inventory* (MMPI).
1946: Se considera a la psicología clínica como profesión y como rama de la psicología.
Stevens propone sus cuatro escalas de medición: nominal, ordinal, intervalo y razón.
1949: Se publica la Escala Wechsler de Inteligencia para Niños (WISC).
1950, México: Se constituye la Sociedad Mexicana de Psicología.
1952: La Asociación Psiquiátrica Americana publica el Manual Diagnóstico y Estadístico (DSM I).
1953: Aparece el primer Código Ético para psicólogos publicado por la *Asociación*

Psicológica Americana (*Ethical principles of psychologist*).
Skinner edita su obra *Ciencia y Conducta Humana.*
Aparición de la Terapia de la conducta.
1954: Primera edición de los estándares técnicos sobre las pruebas (*Technical Recomendations for Psychological Tests and Diagnostic Techniques*).
1957: Edwards presenta sus trabajos sobre la "deseabilidad social" en los cuestionarios de personalidad.
Cronbach y Gleser publican *Psychological Tests and Personal Decision*, donde proponen una teoría formalizada para aplicar en el proceso de evaluación.
1961: Kirk y McCarthy publican la Prueba de Habilidades Psicolingüísticas Illinois.
1965: Kanfer y Saslow, en su trabajo *Behavioral Analysis*, plantean un nuevo modelo de psicodiagnóstico desde la perspectiva conductual.
1967: Anastasi denuncia la enorme separación entre la evaluación psicológica y la psicología en su trabajo *Psychology, Psychologists and Psychological Testing.*
1968: Mischel publica su libro *Personalidad y Evaluación*, que es una dura crítica al modelo de evaluación tradicional de las pruebas.
1969: Bandura publica *Principios de Modificación de Conducta.*
Nancy Bayley edita las Escalas Bayley de Desarrollo Infantil.
1970: Kanfer y Phillips publican *Fundamentos de Aprendizaje en la Terapia del Comportamiento.*
México: El Manual Moderno publica, con normas para México, el *Inventario de Ansiedad Rasgo-Estado* de Spielberger y Díaz-Guerrero.
1972: McCarthy edita sus Escalas McCarthy de Aptitudes y Psicomotricidad para Niños.
1973, México: Se separa la Facultad de Psicología en la UNAM de la Facultad de Filosofía y por primera vez la psicología es reconocida como una profesión, y es por tanto regulada por el Dirección General de Profesiones.
1974: Rudolf Moos publica *The social climate scales: an overview* (Escalas de Clima Social), importante esfuerzo de evaluación del ambiente.
Friedman y Rosenman popularizan el patrón conductual Tipo A, con propensión a cardiopatía.
1980, México: El Manual Moderno publica, con normas para México, el Cuestionario de 16 Factores de Personalidad, de Cattell, Eber y Tatsuoka.
1984, México: Se publica por primera vez en México el *Código Ético del Psicólogo*, por la Sociedad Mexicana de Psicología.
El Manual Moderno publica ,por primera vez con normas para México, la *Escala de Inteligencia Revisada para el nivel Escolar (WISC-RM)*, esto es, la versión mexicana de la prueba Wechsler de inteligencia para niños (WISC).
1985: La *Asociación Psicológica Americana* publica una revisión de los *Standars for Educational and Psychological Testing.*
1992: La *Asociación Psicológica Americana* publica una revisión de los *Ethical Principles of Psychologist and Code of Conduct.*
1994: La *Asociación Psiquiátrica Americana* publica el DSM-IV.

1995, México: El Manual Moderno publica, con normas para México, el *Inventario Multifásico de la Personalidad Minnesota-2 (MMPI-2)*, de Hathaway y Mckinley.
1998, México: El Manual Moderno publica, con normas para México, el *Inventario Multifásico de la Personalidad Minnesota para Adolescente (MMPI-A)*, de Hathaway y Mckinley.
2005, México: El Manual Moderno publica, con normas para México, la prueba de orientación vocacional de Holland, Fritzche y Powell (SDS), *Búsqueda Autodirigida*.
2007, México: El Manual Moderno publica, con normas para México, la *Escala de Inteligencia para el nivel Escolar WISC-IV* de Wechsler.
2008, México: El Manual Moderno publica el DSM-IV.
2011, México: El Manual Moderno publica, con normas para México, la *Escala de Inteligencia para los niveles Preescolar y Primario WIPPSI-III* de Wechsler.
2013, México: El Manual Moderno publica, con normas para México, la *Escala de inteligencia para adultos WAIS-IV*

Una vez descritos los eventos más importantes que dieron origen a la evaluación psicológica como hoy la concebimos, en el capítulo siguiente se abordarán los conceptos fundamentales de esta evaluación.

Referencias

Anastasi, A. (1977). *Tests Psicológicos*. España: Aguilar.

Fernández-Ballesteros, R. (1996). *Introducción a la evaluación psicológica*. Salamanca: Pirámide.

Garaigordobil, L. M. (1998). *Evaluación psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro*. España: Amarú Ediciones.

Garret, H. E. (1975). *Las grandes realizaciones de la psicología experimental*. México: FCE.

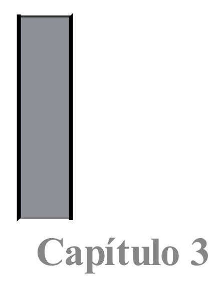
Gregory, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México: El Manual Moderno.

Hogan, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.

Muñiz, J. (2003). *Teoría clásica de los tests*. Madrid: Pirámide.

Nunnally, J. C. (1973). *Introducción a la medición psicológica*. Buenos Aires: Paidós.

Sattler, J. M. (1996). *Evaluación infantil*. México: El Manual Moderno.



Capítulo 3

CONCEPTOS FUNDAMENTALES EN LA EVALUACIÓN PSICOLÓGICA

El objetivo final de toda ciencia es describir, explicar y predecir los fenómenos que estudia; para lograrlo, debe recoger y comparar datos que permitan fundamentar sus teorías. La medida es el componente esencial de este proceso; la medición es lo que le ha permitido a la psicología alcanzar el estatus de ciencia y no sólo ser considerada como una doctrina.

La medición es una actividad prioritaria en la ciencia psicológica en general y en la evaluación psicológica en particular. El proceso de medición, como ya lo mencionábamos, permite a una disciplina alcanzar el estatus de ciencia, ya que, una vez que se pueden medir sus datos, se pueden comparar, replicar, establecer modelos matemáticos para su explicación y realizar investigaciones controladas.

Puesto que la evaluación psicológica objetiva se asienta en la medición, este capítulo, que tiene como meta exponer los conceptos fundamentales en la evaluación psicológica, se iniciará con la definición de medición, su importancia y sus fines; enseguida se describirá una de las técnicas de medición fundamentales del proceso de evaluación: la prueba psicológica; para finalmente llegar a los conceptos de *testing*, psicodiagnóstico y evaluación psicológica.

Importancia de la medición en psicología

La meta de la psicología, como la de cualquier ciencia empírica, es la creación de leyes o principios generales para explicar y predecir los fenómenos que estudia, esto es, pretende la descripción, explicación y predicción del comportamiento de las personas. Para lograr este objetivo, como señala Torgerson (1958, citado en Martínez, 1996), la psicología debe recoger y comparar datos para establecer correlaciones, ecuaciones, etc., que permitan fundamentar las teorías. La medida es un componente esencial de este proceso, porque ha permitido que los modelos matemáticos puedan aplicarse a la ciencia psicológica.

De acuerdo con Nunnally (1973) y Nunnally y Bernstein (1995), el problema más

importante para la psicología es la obtención de métodos adecuados de medición, ya que una teoría sólo puede ser probada si las variables hipotetizadas pueden ser medidas de manera adecuada, es decir, si se pueden establecer relaciones funcionales entre estas variables. Consideran que si una teoría psicológica contiene términos que no pueden medirse adecuadamente (p. ej., atributos postulados como hipótesis, como en la teoría freudiana), será imposible corroborarla, y afirman que los principales progresos en psicología y probablemente en todas las ciencias suceden a un primer desarrollo de los métodos de medición. Como bien mencionan, la psicología no puede progresar más rápido que la medición de sus variables fundamentales.

La psicología estudia variables que no pueden medirse en absoluto, sino sólo aproximadamente, ni directamente, pero sí indirectamente, por ejemplo: inteligencia, ansiedad, atención, depresión, dislexia, etc. Decimos que los atributos psicológicos no pueden medirse absolutamente, porque nunca estaríamos seguros de que hemos medido la totalidad de inteligencia de una persona, toda la ansiedad o depresión que presenta, todos y cada uno de los errores de tipo disléxico que podría cometer en lectoescritura, etc.; por el contrario, en las medidas de atributos físicos, podemos estar seguros de que al medir por ejemplo una mesa, medimos totalmente su longitud y que a esa medida proporcionada por el instrumento no le falta nada, está completa. Por otro lado, la longitud de la mesa la podemos medir directamente, colocamos el metro sobre el atributo de manera directa y así obtenemos la medida que deseamos, a diferencia de la inteligencia, la ansiedad o la depresión, en las que a través de indicadores de la conducta de persona, tanto observables (calificaciones, temblor o sudoración en las manos, aislamiento o inexpresividad) como no observables (sus cogniciones cuando resuelva problemas, su percepción de miedo o catástrofe, su baja autoestima), además de su reporte verbal, podemos inferir la presencia de dicho atributo, pero sólo de manera indirecta. Las experiencias subjetivas, sentimientos, sensaciones y deseos no pueden ser observados directamente y por lo tanto no pueden someterse a medición, pero cuando un sujeto formula un juicio, enuncia una preferencia o simplemente habla con otros acerca de sus experiencias subjetivas, esta conducta o reporte verbal satisface los requisitos de la investigación científica y la medición se vuelve posible.

La ventaja principal de la medición es que erradica conjeturas sobre lo que observamos o escuchamos. Por ejemplo, si como psicólogos atendemos a un paciente y, en función de lo observado y sus respuestas en la entrevista, planteamos la hipótesis de que tiene depresión, en primer lugar tendríamos que verificar si en verdad la tiene —que no es fatiga, ansiedad, tristeza, falta de alegría o cualquier otro problema— y una vez verificado que es depresión, nos interesaría saber qué tanta depresión tiene o en qué grado: leve o moderada, en suficiente cantidad para ameritar tratamiento, grave como para estar en riesgo de suicidio, además, nos interesaría saber también cómo se manifiesta su depresión: su estado de ánimo, problemas de socialización, baja autoestima, preocupaciones, sentimientos de culpa, pérdida de energía, problemas de sueño y apetito, ideación suicida, etc.; lo primero es establecer un diagnóstico, lo segundo, una clasificación. Si no tenemos un instrumento que nos ayude a medir depresión, no

podemos contestar las cuestiones planteadas; como menciona Meehl (1954, citado en Miller, 1983), una vez que se ha planteado una hipótesis, hay que someterla a los cánones habituales de la inferencia y ponerla a prueba mediante los criterios científicos usuales de la evidencia y la probabilidad y, para poder hacer esto, necesitamos un instrumento de medición. Del mismo modo, un instrumento de tal tipo nos ayudaría no sólo a identificar la depresión y clasificarla, sino también a tomar una decisión sobre la intervención más adecuada y, terminado el tratamiento, corroborar que efectivamente el problema fue eliminado o significativamente aminorado, es decir, verificar la predicción y comprobar el cambio.

¿Qué medimos en psicología?

La respuesta más sencilla a esta pregunta es: lo que medimos en psicología es la conducta, el comportamiento. Pero esta respuesta no nos dejaría satisfechos, ya que, por ejemplo, no tiene el mismo nivel de dificultad medir una conducta como "fumar" o "hacer berrinche" (que son conductas observables), que conductas tales como "nerviosismo", "estudiar", "estar triste", "bajo rendimiento" o "problema de aprendizaje". En realidad, no podemos medirlas de manera directa ni absoluta, entonces, más que conductas, lo que medimos son "atributos de la conducta", más precisamente, "atributos psicológicos". Por otra parte, aunque una conducta como "fumar" sea fácilmente medible, a la psicología le interesaría primordialmente el componente psicológico de la conducta de fumar, ya que ésta tiene otros componentes: biológico, fisiológico o motor; así, de la conducta de fumar, nos interesaría el componente psicológico o atributo psicológico, por ejemplo: la dependencia psicológica hacia el cigarrillo, las cognitivas que el sujeto presenta tanto cuando se va a fumar el cigarro como cuando no puede hacerlo, otras conductas que desencadenan la conducta de fumar, las consecuencias tanto agradables como aversivas de fumar, etc.

Para que quede un poco más clara la distinción entre medir un objeto o un atributo, digamos que cuando medimos una mesa (un objeto físico), más que la mesa como tal, medimos ciertos atributos de ella: largo, ancho, altura; del mismo modo, cuando medimos personas o comportamientos, no lo hacemos absolutamente, sino sólo ciertos atributos psicológicos que nos interesan: motivación, ansiedad, depresión, inteligencia, afrontamiento, personalidad, bajo rendimiento, problema de aprendizaje, dislexia, etc.

Sin embargo, a diferencia de la mesa —cuyos atributos largo, ancho o altura los podemos medir directa y absolutamente por medio de un instrumento que mida longitud —, muchos de los atributos psicológicos no los podemos medir ni directa ni absolutamente, sino sólo indirectamente y de manera aproximada, por eso se habla de que los atributos psicológicos se miden mediante indicadores, esto es, instrumentos que nos van a mostrar indirectamente una muestra del atributo que pretendemos medir. Por ejemplo, si intentáramos medir el problema de aprendizaje de un niño catalogado como disléxico, no podríamos medir la dislexia directamente, pero sí una muestra de la clase y el número de errores disléxicos que presenta en lectura, copia y dictado (omisiones,

inserciones, errores de secuenciación, confusión por la dirección de los rasgos que forman las letras, etc.), y las consecuencias que estos errores tienen en su aprovechamiento escolar, en sus conductas motoras (agresividad, hiperactividad), en sus emociones y cogniciones (baja autoestima).

¿Qué es la medición?

El objetivo de la medición en psicología es determinar cuánto está presente de un atributo en determinada persona, y esto requiere una expresión numérica de la cantidad. En el nivel más bajo de cuantificación, la medición se ocupa de la presencia o ausencia del atributo; en el nivel intermedio, determina si se posee más o menos cantidad del atributo medido, y, en el nivel más alto, qué tanto se posee del atributo y si se pueden cuantificar las diferencias presentadas en él.

Para Lord y Novick (1968, citado en Martínez, 1996, p. 17), la medición es "…un procedimiento para la asignación de números (puntuaciones, medidas) a propiedades específicas de unidades experimentales, de modo que las caractericen y preserven las relaciones especificadas en el dominio comportamental"; según Stevens (1951, *op. cit.*, p. 22), medida es "…la asignación de números a objetos o sucesos según ciertas reglas"; para Nunnally (1973, p. 23), la medición "…consiste en un conjunto de normas para asignar números a los objetos de modo tal que estos números representen cantidades de atributos".

De esta manera, la medición se ocupa del mundo real en términos de **propósitos, operaciones y validez**. El propósito es cuantificar y clasificar los atributos de las personas y objetos reales para definir si las propiedades medidas caen en las mismas o en diferentes categorías; las operaciones tienen como objetivo obtener medidas de acuerdo con un conjunto de normas, y la validez o utilidad de una medida depende siempre del carácter de los datos empíricos (Nunnally, 1973; Nunnally y Bernstein, 1995).

En el apartado anterior vimos que lo que medimos no son las personas ni los objetos, sino ciertos atributos de ellos; así, el término **atributo** se refiere a características determinadas y particulares de los objetos. Al respecto, Nunnally (1973) y Nunnally y Bernstein (1995) mencionan que hay que considerar cuidadosamente la naturaleza del atributo antes de medirlo, ya que es posible que no exista, o al menos no en la forma propuesta, y puesto que la medición en psicología requiere de un proceso de abstracción, es probable que el atributo sí exista, pero esté mal medido o mal definido.

En las definiciones se destaca que se emplean **números** para representar cantidades. La cuantificación se refiere entonces a la asignación de números para informar qué cantidad de un atributo está presente en un objeto; la cuantificación está tan estrechamente ligada al concepto de medición que los dos términos se intercambian con frecuencia.

Los procedimientos para asignar números a los atributos medidos deben formularse explícitamente. De esta manera, cualquier conjunto de **normas** que cuantifique sin ambigüedad las propiedades de los objetos constituye una medida legítima y adquiere el

derecho de competir con otras en cuanto a utilidad científica (Nunnally, 1973).

Para Nunnally y Bernstein (1995), las normas o reglas son un aspecto importante de la estandarización*. Una medida está estandarizada si: 1) sus reglas son claras, 2) su aplicación es práctica, 3) no requiere de una gran habilidad de los administradores para llevarla a cabo y 4) sus resultados no dependen del administrador específico; sin embargo, también mencionan que hay que considerar que la claridad de la medición no necesariamente garantiza el poder explicativo.

El punto central de la estandarización es que los usuarios de un instrumento determinado deben obtener resultados similares; por ejemplo, podemos decir que una determinada prueba está bien estandarizada si diferentes examinadores obtienen puntajes similares al evaluar a un niño en particular en un momento determinado.

Los métodos estandarizados proporcionan resultados numéricos más detallados que los juicios personales, esto es, la comunicación es más sencilla cuando se dispone de medidas estandarizadas.

Finalmente, la **utilidad** de un método de medición se establece en función del grado en que los datos obtenidos: 1) correspondan a un modelo matemático, 2) midan un atributo individual, 3) puedan ser repetibles bajo circunstancias similares, 4) sean válidos en diversos sentidos y 5) produzcan relaciones interesantes con otras medidas científicas (Nunnally y Bernstein, 1995).

Podemos resumir que las medidas estandarizadas son objetivas, ya que nos permiten eliminar las conjeturas sobre lo que observamos o escuchamos, esto es, sobre la observación no científica (un principio clave de la ciencia es que cualquier exposición de hechos realizada por un científico debe ser verificable de manera independiente por otros científicos), facilitar la comunicación, ahorrar tiempo y ayudar a la generalización científica.

De esta manera, Martínez (1996) establece que la medida comienza con un procedimiento para identificar los elementos del mundo real con los elementos o constructos del sistema lógico, por medio de una definición semántica precisa; una vez hecho esto, se procede a la medición propiamente dicha, para lo cual deben seguirse los siguientes pasos: 1) identificar el objeto o sujeto que se medirá, 2) identificar la propiedad del objeto o de la conducta que se quiere medir en el sujeto (atributo), 3) seleccionar el instrumento y 4) aplicar una regla de asignación numérica por la que se otorga una cifra a la propiedad de la unidad experimental (objeto o sujeto) que se mide. Cabe señalar que para medir la propiedad o el atributo, debe existir isomorfismo entre las características del sistema numérico y las relaciones entre las diversas cantidades de la propiedad medida. Por ejemplo, si después de una entrevista en una sesión de psicología clínica sospechamos o hipotetizamos, por los síntomas que la persona nos refiere, que presenta ansiedad (aquí estamos realizando los pasos 1 y 2), enseguida seleccionamos un instrumento que mida dicho atributo, por ejemplo el Inventario de Situaciones y Respuestas de Ansiedad de Tobal y Cano (paso 3), y después de aplicarlo obtenemos cuatro medidas de ansiedad en percentiles: ansiedad total, ansiedad cognitiva, ansiedad fisiológica y ansiedad motora (paso 4). En este instrumento psicométrico, al igual que

todos los que están publicados, existe isomorfismo entre las cantidades obtenidas en su aplicación y la manifestación del atributo medido, en este caso, a mayor puntuación, mayor ansiedad y viceversa.

Para hablar de isomorfismo, tenemos que remitirnos al concepto de escalamiento. El escalamiento es el desarrollo de reglas sistemáticas y de unidades significativas de medida que nos permiten cuantificar las observaciones empíricas; de esta manera, se establece una escala de medición cuando se define el conjunto de valores posibles que pueden asignarse y, lo más importante, se determina la regla de asignación que establece la correspondencia entre el sistema empírico (lo que vamos a medir) y el sistema numérico (Martínez, 1996).

La teoría del escalamiento o teoría representacional de la medición de Stevens (1951, citado en Martínez, 1996) gira en torno a tres grandes áreas: representación, unicidad y significación.

1. **El problema de la representación.** La medición debe encontrar un sistema relacional numérico, con una estructura semejante al sistema relacional empírico que se pretende medir, es decir, el sistema relacional numérico puede representar al sistema relacional empírico porque hay isomorfismo entre las propiedades medidas en el segundo con las propiedades numéricas del primero.
   El sistema relacional numérico está conformado por las escalas de medición nominal, ordinal, de intervalo y de razón. La escala nominal es el nivel más elemental de medición, agrupa las características que se medirán en subconjuntos o clases (el atributo puede estar o no presente) y sólo usa los números para clasificar los objetos (1: hombre; 2: mujer). En la escala ordinal, observamos diferentes grados de un atributo y se establece una relación de mayor y menor cantidad de esa propiedad (puntuaciones de depresión: sujeto A centil 80; sujeto B centil 60; sujeto C centil 40; podemos entonces decir que A>B>C en cuanto a depresión, pero no que la diferencia entre la depresión de A y B sea la misma que entre la de B y C, ni que A tenga el doble de depresión que C). Las escalas de intervalo y razón indican no sólo un orden decreciente o creciente en la magnitud del atributo medido, sino que las diferencias de ésta son iguales a los intervalos entre los números asignados a dichas características, sin embargo,en la escala de intervalo no existe un cero absoluto, es decir, el numeral cero no indica ausencia de la característica medida; un ejemplo de escala de intervalo es la medición de la inteligencia, donde si un individuo obtuviera un CI de 0 esto no significaría ausencia de inteligencia. Cuando sí existe un cero absoluto, esto es, ausencia de la característica medida, la escala se llama de razón. Por ejemplo, una persona pesa 100 kg al iniciar un tratamiento de reducción de peso, un año después pesa 75 kg y dos años más tarde, 50 kg; la diferencia entre el primer y el segundo año es igual a la que se dio entre el segundo y el tercero: 25 kilogramos, al iniciar el tratamiento pesaba el doble que al finalizar. Cero kg indica ausencia de la característica, como podría haber sido si de una medición a otra hubiera conservado el mismo peso, esto es, hubiera bajado cero kg.

2. **El problema de la unicidad** se refiere a que los números elegidos arbitrariamente para ser asignados en la medición pueden cambiarse por otros, siempre y cuando no se modifiquen las relaciones numéricas que representan las relaciones empíricas (p. ej., 1: hombre, 2: mujer, no cambia la información si se cambia por: 1: mujer, 2: hombre).
3. **El problema de la significación** tiene que ver con la validez de una conclusión empírica inferida a partir de una conclusión numérica es decir, si la propiedad a medir existe o no existe, si se presenta mayor o menor cantidad del atributo y si se pueden cuantificar esas diferencias que se presentan; plantea además los estadísticos admisibles para cada tipo de escala (p. ej., los números de la escala nominal no se pueden sumar, ni tampoco los de la escala ordinal, pero sí los de las escalas de intervalo y de razón).

Adentrémonos ahora en una de las técnicas de medición más usadas por la psicología: las pruebas psicológicas, las cuales, como ya vimos en los capítulos anteriores, se empezaron a gestar con Galton, Cattell les dio nombre y apellido (test mental) y Binet elaboró la primera prueba considerada como psicométrica. Estos autores contribuyeron, conjuntamente con la psicología experimental, a echar por tierra la predicción de Kant de que (citado en Nunnally, 1973, p. 20): "...la psicología nunca será una ciencia porque sus datos básicos no pueden observarse ni medirse".

Pruebas psicológicas

Las pruebas psicológicas son sólo una técnica, entre muchas, de medición en psicología; les dedicamos un apartado porque precisamente los dos últimos capítulos de esta obra están dirigidos a exponer su selección, aplicación, calificación e interpretación.

Test es una palabra inglesa que significa prueba, reactivo, ensayar, probar o comprobar; este término tiene su origen a su vez en la raíz latina *testis*, que figura en palabras como testigo y testimonio. La suposición implícita desde el punto de vista psicométrico es que las pruebas miden las diferencias individuales en cuanto a algún rasgo o atributo —si suponemos que todas las personas poseen la característica medida en diferentes cantidades— y su propósito es estimar esa cantidad.

Se han propuesto varias definiciones de la palabra *test*, que nosotros traducimos en nuestro idioma como "prueba"; algunas de ellas son:

Para Anastasi (1977, p. 21), una prueba psicológica constituye esencialmente una medida objetiva y tipificada de una muestra de conducta.

Brown (1980, p. 6) define una prueba como un procedimiento sistemático para medir una muestra de conducta.

De acuerdo con Fernández-Ballesteros (1996, p. 139), la prueba es un instrumento sistemático y tipificado que permite la comparación de un sujeto con un grupo normativo.

Para Hogan (2004, p. 30), una prueba es un proceso o instrumento estandarizado que

genera, en forma cuantificada, información sobre una muestra de comportamiento o proceso cognitivo.

Gregory (2001, p. 36) menciona que una prueba es un procedimiento estandarizado para tomar una muestra de conducta y describirla en categorías o puntuaciones, la mayor parte con normas o estándares, con base en los cuales pueden utilizarse resultados para pronosticar otras conductas más importantes.

Para Yela (1980, citado en Martínez, 1996, p. 32), una prueba es una situación problemática previamente dispuesta y estudiada a la que el sujeto ha de responder bajo ciertas instrucciones y de cuyas respuestas se estima, en comparación con las respuestas de un grupo normativo (o un criterio u objetivo), la calidad, índole o grado de algún aspecto de su personalidad.

Garaigordobil (1998, p. 23) dice que "prueba" se refiere a una situación controlada, en la que se recogen muestras de conducta de sujetos que responden a ciertos estímulos. Estas respuestas son puntuadas o valoradas según algunos criterios y ofrecen información del lugar que ocupa el sujeto dentro de un grupo de referencia normativo.

Estas definiciones nos ayudarán a desglosar las principales características que debe poseer una prueba psicológica: medida objetiva, muestra de conductas, técnica sistemática, establecimiento de normas de comparación y predicción o inferencia (Martínez, 1996; Garaigordobil, 1998; Gregory, 2001; Hogan, 2004).

Medida objetiva. Implica precisión en la medida; hace posible que dos evaluadores independientes que siguen las instrucciones de la prueba tengan el mismo resultado. Así pues, la aplicación, la puntuación y la interpretación de las puntuaciones son objetivas en cuanto son independientes del juicio subjetivo de un examinador en particular.

Muestra de conductas. Las pruebas no pueden medir de manera absoluta un atributo, sino sólo aproximadamente, son una muestra limitada de las conductas que representan el dominio conductual del atributo que se medirá. Una prueba es una muestra de todos los reactivos posibles a partir de las cuales se establecen predicciones en función de que se emplea una muestra representativa del atributo de interés. En este sentido, la muestra de conducta será útil sólo cuando permita al examinador realizar inferencias acerca del dominio total de conductas relacionadas.

Técnica o procedimiento sistemático. Se refiere a la noción de estandarización o tipificación. El procedimiento estandarizado es una característica esencial de cualquier prueba psicológica y supone uniformidad de procedimientos en la aplicación y puntuación de la misma. La prueba siempre se aplica a los diferentes sujetos bajo las mismas condiciones establecidas en el manual; así, una prueba se construye, se administra y califica según reglas preestablecidas. Si van a compararse las puntuaciones obtenidas por varios sujetos, las condiciones de aplicación de la prueba deben ser las mismas para todos. La tipificación o uniformidad se extiende a los materiales empleados, los límites de tiempo, las instrucciones orales a los sujetos, las demostraciones previas, las formas de resolver las dudas planteadas y todos los demás detalles de la situación de la prueba. Se considera que una prueba está estandarizada si los procedimientos para su aplicación son uniformes de un examinador a otro y de un ambiente a otro.

Otro paso importante en la tipificación de una prueba es el **establecimiento de normas,** ya que sin ellas es imposible interpretar las puntuaciones obtenidas. Las pruebas psicológicas no tienen modelos predeterminados de clasificación de las puntuaciones, por lo que la puntuación de un individuo sólo puede valorarse cuando se compara con las de otros. Como su nombre lo indica, una norma es la actuación media o normativa, así pues, las pruebas psicológicas se basan en normas establecidas empíricamente: las respuestas dadas por un sujeto a una prueba se comparan con las de un grupo normativo, con el propósito de clasificar su actuación en el atributo medido con respecto al grupo de comparación. De esta manera, una prueba psicológica permite la derivación de puntuaciones o categorías. Thorndike (1918, citado en Gregory, 2001,) expresó el axioma esencial de las pruebas: "Aquello que existe de alguna manera, existe en cierta cantidad". McCall (1939, citado en Gregory, 2001) fue un paso más allá al declarar: "Cualquier cosa que existe en cierta cantidad puede medirse". Toda prueba suministra una o más puntuaciones o proporciona evidencia de que una persona pertenece a una u otra categoría. En pocas palabras, las pruebas psicológicas resumen la ejecución de las personas en números o clasificaciones.

Predicción o inferencia. La aplicación de una prueba implica hacer inferencias o predicciones futuras fundamentadas en la ejecución observada del sujeto en la prueba.

Una vez que hemos clarificado los conceptos sobre medición, medición psicológica y pruebas, podemos finalmente abordar tres conceptos clave en la evaluación psicológica: testing, psicodiagnóstico y evaluación psicológica.

Testing

Como ya vimos en el capítulo 2, una de las consecuencias de las dos guerras mundiales fue la proliferación de pruebas psicológicas, en ninguna otra época se había llevado a cabo un programa tan grande de construcción de pruebas. A decir de Pelechano (1988, citado en Garaigordobil, 1998), surge entonces un grupo de psicólogos que argumentaba que para a aceptar algo como científico debería expresarse en números, a la vez que otro grupo de psicólogos rechazaba radicalmente todo lo que significara expresión numérica. Fernández-Ballesteros (1996) menciona que a esta disciplina de expresarlo todo con números se le llegó incluso a denominar testing, esto es, testing se refiere a la mera aplicación de pruebas con el propósito de adjudicarle un número o asignarle una clasificación al individuo al que se le está aplicando la prueba; la aplicación de pruebas no como un medio de obtener más información, sino como el fin último del proceso.

De esta manera, testing se refiere únicamente a un procedimiento mecánico, tipificado, cuantitativo y objetivo, mientras que evaluación es un proceso complejo que comprende desde el planteamiento del problema hasta la emisión de un juicio o toma de decisiones; la evaluación psicológica no sólo se limita al uso de pruebas como técnica de medida, sino que utiliza también otras técnicas como la entrevista, observaciones, registros psicofisiológicos, etc. Por lo tanto, testing no es lo mismo que evaluación; la mera aplicación de pruebas no tiene sentido, no es un fin sino un medio; ningún psicólogo

aplica una prueba sólo por el gusto de aplicarla y obtener una medida de algo, la aplicación de pruebas es sólo una técnica justificada para obtener una medida de un atributo relevante dentro de un proceso evaluativo. Como bien señala Garaigordobil (1998), la crítica al testing de medir por medir y a las evaluaciones psicométricas sin objetivos específicos surge cuando no se tiene un conocimiento amplio de un marco de relaciones en donde se inserte la medida, lo cual la deja sin valor.

Psicodiagnóstico

Históricamente, el término psicodiagnóstico, como vimos en el capítulo 2, fue utilizado por primera vez en 1921 por Rorschach, quien tituló así su obra donde presenta la Prueba de las Manchas de Tinta (*Psychodiagnostik*). Etimológicamente, diagnóstico proviene del griego *dia-gignosko*, que a su vez procede del latín *gnoscere*, y significa aprender a conocer, tener conocimiento, entender algo, conocer a fondo o en profundidad. El *Diccionario de la Real Academia Española* (2001) subraya el carácter médico del término diagnóstico y lo define como el arte o acto de conocer la naturaleza de una enfermedad mediante la observación de sus síntomas y signos.

En psicología, el concepto de psicodiagnóstico, quizá por su herencia médica, se ha asociado prioritariamente a la clínica y al estudio de las patologías psicológicas o enfermedades mentales:

> La definición dada del psicodiagnóstico en su acepción tradicional hace referencia al sentido médico-psiquiátrico del término, ya que se plantea como objetivo exclusivamente la clasificación taxonómica. Es un diagnóstico dirigido a la exploración de aspectos psicopatológicos, negativos o deficitarios del funcionamiento psicológico, y la aplicación de instrumentos tiene por finalidad complementar, corroborar y medir aquellas tendencias que, según la experiencia práctica de los clínicos, son de mayor significación. (Garaigordobil, 1998, p. 25).

Para Pelechano (1988, citado en Garaigordobil, 1998), el término psicodiagnóstico tiene claras connotaciones médico-psiquiátricas:

- La conducta observable no posee una entidad que se agote a sí misma como síntoma de algo
- Predomina lo cualitativo frente a lo cuantitativo, un tipo de conocimiento situado entre la opinión y la ciencia, basado principalmente en la experiencia del diagnosticador
- Se utiliza un método que consiste básicamente en una interacción verbal y no verbal entre el psicodiagnosticado y el psicodiagnosticador
- Se pretende no sólo estimar el estado actual del sujeto (se enfoca sobre lo patológico), sino que también intenta formular un juicio pronóstico
- Predomina la observación y el registro para identificar la alteración psicológica y posteriormente llevar a cabo el tratamiento

Sin embargo, si bien sus comienzos denotaban la identificación de alteraciones psicopatológicas, en la actualidad el término sigue vigente y se le usa tanto como sinónimo de evaluación psicológica como para identificar el problema psicológico del sujeto evaluado, o bien, su clasificación. En este libro los vamos a considerar como sinónimos (a menos que se especifique que únicamente nos interesa la detección y/o clasificación), aunque preponderantemente utilizaremos el de evaluación psicológica.

Evaluación psicológica

De acuerdo con Garaigordobil (1998), el término inglés *assessment* aparece por primera vez en 1948, en el libro *Assessment of Men*, donde se describen las actividades realizadas por la Sección de Psicología de la Oficina de Servicios Estratégicos de los Estados Unidos, encabezada por Murray; en esa obra se utiliza el término de evaluación en vez de psicodiagnóstico, pues se deja de lado la detección de aspectos patológicos, y en su lugar se enfatizan los aspectos positivos y de desarrollo potencial del ser humano.

Otro momento del auge del término fue durante los años sesenta, con la aparición, dentro de la psicología conductual, de la disciplina de la evaluación conductual como una alternativa al diagnóstico nosológico psiquiátrico y como respuesta a las necesidades surgidas a partir de la modificación y terapia de la conducta, que enfatizaba que debería haber una estrecha relación entre la evaluación y el tratamiento, ya que la información obtenida en la evaluación permitiría tomar decisiones encaminadas al cambio o la modificación de la conducta (evaluar para intervenir y evaluar a su vez el tratamiento). Así, se definía a la evaluación conductual como la alternativa a la evaluación psicológica, cuya meta es la identificación de las conductas objeto de estudio motoras, fisiológicas y cognitivas, así como de las variables ambientales y/o personales que las mantienen o controlan, con el objetivo último de planear la intervención (Fernández-Ballesteros, 1994).

En un sentido estrecho, la evaluación en psicología consiste en diagnosticar, es decir, detectar la presencia de un estado, un conocimiento, una patología, etc., de acuerdo con el campo donde se esté evaluando. En un sentido amplio, el término evaluación no se restringe sólo al diagnóstico, sino a todo un proceso cuya meta consiste en identificar, rotular, seleccionar, plantear objetivos, intervenir, pronosticar y evaluar si los objetivos planteados se alcanzaron; su finalidad estriba sobre todo en su utilidad para la toma de decisiones.

Así, la evaluación psicológica no se conforma únicamente con aplicar algún tipo de prueba o instrumento para conocer el estado actual de un sujeto o grupo de sujetos en cuanto a algún atributo o problema psicológico, sino que va más allá, y éste sólo sería el primer paso de un proceso que involucra una toma de decisiones que comprende, a su vez, seleccionar y plantear objetivos de terapia, seleccionar las técnicas y estrategias para la intervención, evaluar continuamente para constatar que los objetivos planteados se van alcanzando, así como para detectar posibles problemas en su consecución y evaluar para conocer si las metas finales se alcanzaron. De este modo, la evaluación no representa un

fin en sí misma, más bien persigue un objetivo eminentemente práctico y como tal está al servicio de la disciplina: su objetivo fundamental no es la explicación, sino la búsqueda de medidas prácticas de intervención.

La evaluación psicológica entraña un concepto muy importante: las inferencias, basadas en las respuestas que se dan en la situación de evaluación. Como mínimo, el proceso de evaluación supone que la conducta que se evaluará mediante algún instrumento específico es una muestra de los repertorios de conducta que el evaluado posee; representa lo que el individuo ha aprendido como resultado de su interacción con el entorno social.

Sin embargo, así como no todas las evaluaciones son iguales, tampoco lo son sus fines. Al respecto, podemos considerar que dentro del proceso de evaluación psicológica existen estrategias selectivas y estrategias modificativas. Dentro de las primeras se busca la optimación mediante la selección adecuada de personas —en determinadas condiciones, se debe elegir a los individuos con un criterio de ejecución o clasificarlos de acuerdo con dicho criterio— o de condiciones —se parte de unos individuos concretos y la estrategia consiste en buscar las condiciones idóneas que mejor se adapten a cada uno de ellos según el criterio de optimación. Un ejemplo representativo de este tipo de estrategias es el que se realiza en psicología organizacional, ya sea buscando a las personas que cumplan el “perfil” de un determinado puesto, o bien, instalar a las personas en un trabajo específico según sus habilidades, competencias y aptitudes. Por otra parte, en la estrategia modificativa se busca la optimación por medio de la modificación de las conductas de las personas y de las condiciones. Esta estrategia es eminentemente clínica y educativa; una vez evaluada la persona, los objetivos irán encaminados a revertir, minimizar, modificar o eliminar el problema detectado. Dentro de esta última estrategia se debe considerar que para que sea útil debe proporcionar los elementos suficientes para determinar el método modificativo en cada caso concreto y así verificar la efectividad del mismo una vez utilizado (Pawlik, 1980).

El objetivo final que persigue toda evaluación es la toma de decisiones, la cual presupone la existencia de alternativas entre las cuales hay que elegir; así, se podría considerar entonces a la intervención psicológica como todo aquello que al final de un proceso evaluativo se puede implementar.

En las tres últimas décadas, el término de evaluación psicológica ha irrumpido con mucha fuerza dentro del campo de la psicología, y ha dejado de lado tanto las concepciones originales del psicodiagnóstico —con sus claras connotaciones médico-psiquiátricas dirigidas fundamentalmente a la detección de aspectos psicopatológicos o deficitiarios del funcionamiento psicológico— como las reduccionistas de la evaluación conductual, que si bien intentaron identificar conductas objeto de estudio motoras, fisiológicas y cognitivas, así como las variables ambientales y personales que las mantienen y controlan, en muchas ocasiones sus planteamientos radicales pretendían tener la exclusividad de la evaluación y rechazar cualquier otra. A este respecto, y ya que la evaluación psicológica es un proceso complejo que va desde el planteamiento del problema hasta la recomendación de soluciones y la toma de decisiones, estamos

completamente de acuerdo con Silva cuando señala que ninguna perspectiva, estrategia, procedimiento o técnica tiene la exclusiva ((1983, citado en Garaigordobil, 1998).

En el proceso de evaluación psicológica se distinguen dos fases: la primera es la identificación de los atributos que se medirán de acuerdo con el objetivo planteado por la evaluación, la segunda es la selección y el diseño de técnicas confiables y válidas para medir y evaluar esos atributos; su objetivo primordial es el análisis científico de la conducta humana para poder llegar a su descripción, nosología, comprensión, análisis, explicación y predicción. Para Pelechano (1988, citado en Garaigordobil, 1998), la evaluación psicológica es un concepto multidimensional, ya que se puede evaluar más de un atributo si así se considera necesario (multirrasgo), utilizar para tal fin más de un método de obtención de información (multimétodo), recoger información de más de una fuente (multifuente) y participar a veces más de un evaluador (multijueces). Por ejemplo, a un niño con bajo rendimiento académico y con problemas de socialización, se puede considerar necesario aplicarle una prueba de inteligencia, otra para medir sus habilidades de interacción social y otra para descartar ansiedad; además, se puede contemplar realizar una entrevista con el propio niño, revisar sus cuadernos, y pedirle a otro evaluador que lleve a cabo observaciones en vivo en el salón de clases y a la hora del recreo para tener más datos acerca de su interacción con compañeros; entrevistar también a su profesor e incluso pedirle que nos conteste una prueba que califique la socialización del niño.

Por su parte, Forns (1985, citado en Garaigordobil, 1998) plantea algunas reflexiones importantes sobre la evaluación psicológica. Considera al sujeto evaluado como un ser biológico y sociocultural —que se mueve en diversos ambientes que condicionan y reactivan su conducta—, activo, interactuante y normalmente capaz de autorregulación. En este caso, el proceso de evaluación psicológica se encaminaría hacia el logro de un mejor desempeño autorregulador.

Llegamos por fin a la definición de evaluación psicológica que nos parece más completa, la planteada por Fernández-Ballesteros (1996):

La evaluación psicológica es una disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento de un sujeto o un grupo de sujetos —a los niveles de complejidad que se estime oportunos, como podrían ser el motor, el fisiológico y/o el cognitivo-, con distintos objetivos aplicados y de investigación (detección, diagnóstico, descripción, clasificación, selección, explicación, predicción, tratamiento o intervención, valoración), por medio de la aplicación de diferentes dispositivos, pruebas y técnicas de medición y evaluación, cuya finalidad última es la toma de decisiones.

Actualmente, se considera a la evaluación psicológica tanto una rama de la psicología como una disciplina en sí misma, con objetivos básicos y de investigación, dedicada al análisis científico de la conducta humana en los diferentes niveles de complejidad y contextos en que se desarrolla, cuya meta es la toma de decisiones para la intervención. En esta definición, como lo menciona Silva (1983, *op. cit.*), ninguna perspectiva, estrategia, procedimiento o técnica, tiene la exclusiva.

En evaluación psicológica, para poder utilizar las pruebas psicológicas de manera

acertada, eficiente y ética, debemos conocer los fundamentos psicométricos que las sustentan, tanto para seleccionar las pruebas que vamos a utilizar en la evaluación de una persona según el propósito que persigamos con su aplicación como para poder interpretar adecuadamente los resultados de su ejecución, lo cual es el tema del siguiente capítulo.

Referencias

Anastasi, A. (1977). *Tests psicológicos*. España: Aguilar.

Brown, F. G. (1980). *Principios de la medición en psicología y educación*. México: El Manual Moderno.

Diccionario de la Real Academia de la Lengua Española. (2001). Recuperado de http://buscon.rae.es/draeI/

Fernández-Ballesteros, R. (1994). *Evaluación conductual hoy. Un enfoque para el cambio en psicología clínica y de la salud*. Madrid: Pirámide.

Fernández-Ballesteros, R. (1996). *Introducción a la evaluación psicológica*. Salamanca: Pirámide.

Garaigordobil, L. M. (1998). *Evaluación psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro*. España: Amarú Ediciones.

Gregory, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México: El Manual Moderno.

Hogan, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.

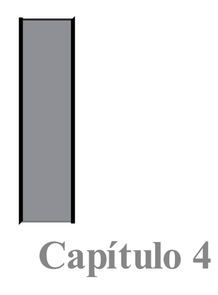
Martínez A. R. (1996). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Síntesis.

Miller, G. A. (1983). *Introducción a la psicología*. Madrid: Alianza Editorial.

Nunnally, J. C. (1973). *Introducción a la medición psicológica*. Buenos Aires: Paidós.

Nunnally, J. C. y Bernstein, I. J. (1995). *Teoría psicométrica*. México: McGraw-Hill.

Pawlik, K. (1980). *Diagnosis del diagnóstico*. Barcelona: Herder.



Capítulo 4

FUNDAMENTOS PSICOMÉTRICOS EN LA EVALUACIÓN PSICOLÓGICA

Así como en las ciencias físicas la elaboración de instrumentos se deriva de la teoría en la que están inmersos y de una tecnología que permite medir sus objetos de estudio, los instrumentos de evaluación psicológica responden a una teoría de la medida; como menciona Martínez (1996), los instrumentos de medición psicológica necesitan demostrar su utilidad y validez científica. En psicología, la psicometría es la que garantiza que los instrumentos de medición elaborados cuenten con las garantías científicas para su uso.

La psicometría es el conjunto de métodos, técnicas y teorías implicados en la medición de variables psicológicas; estudia las propiedades métricas exigibles en las mediciones psicológicas y establece las bases para que éstas se realicen de manera adecuada. El objetivo de la psicometría es proporcionar modelos para transformar los hechos en datos con la finalidad de asignarles valores numéricos a los sujetos, sobre la base de sus respuestas (Muñiz, 2003).

En este capítulo estudiaremos cuáles son los fundamentos psicométricos que caracterizan a la evaluación psicológica; para tal fin, presentaremos la definición de psicometría y una breve reseña histórica de su nacimiento; abordaremos la teoría clásica de las pruebas; expondremos las garantías científicas que toda prueba debe poseer: confiabilidad y validez; intentaremos una clasificación de las pruebas y, finalmente, explicaremos cuáles son los diferentes tipos de puntuaciones que nos proporcionan las pruebas y su interpretación.

Definición y reseña histórica

En general, la psicometría trata de todo aquello relacionado con la medición psicológica; en particular, y lo que la hace diferente de otras disciplinas psicológicas, es que se especializa en las propiedades métricas que dichas mediciones deben poseer, independientemente del campo de aplicación y de los instrumentos utilizados; su objetivo es proporcionar modelos matemáticos para poder transformar los hechos en datos, con la

finalidad de poderles asignar números a las respuestas dadas por los sujetos en la evaluación realizada.

Para Muñiz (2003), el nacimiento de la teoría de las pruebas se origina con los trabajos de Spearman, en 1904, acerca de su teoría de los dos factores de la inteligencia, y en 1907, cuando acuña conjuntamente con Krueger el término coeficiente de confiabilidad, para establecer así los fundamentos de la teoría de las pruebas. El objetivo central que Spearman perseguía era encontrar un modelo estadístico que explicase adecuadamente las puntuaciones obtenidas en las pruebas y permitiera la estimación de los errores de medida implícitos en todo proceso de medición. También, en 1904, Thorndike publica el libro *Introduction to the Theory of Mental and Social Measurements* (Introducción a la teoría de la medición mental y la medición social).

A partir de esas publicaciones, se desarrolló una creciente actividad en el campo de la psicometría, tanto en la teoría como en la construcción y tecnología de las pruebas (Muñiz, 2003):

- En 1931, Thurstone publica su libro "*The Realibility and Valididy of Tests* (Confiabilidad y validez de las pruebas).
- En 1936, se funda la Sociedad Psicométrica Americana, bajo la dirección del mismo Thurstone, quien crea también la revista que publicará a partir de esa fecha todo lo relacionado con la psicometría: *Psychometrika*; asimismo, en ese año, Guilford publica su libro *Psychometric Methods* (Métodos psicométricos).
- En 1946, Stevens da a conocer su trabajo sobre las Escalas de Medición (nominal, ordinal, de intervalo y de razón), sus propiedades y relaciones, así como sus implicaciones en el uso de las distintas técnicas estadísticas.
- En 1947, Thurstone publica su obra *Análisis Factorial Múltiple*, en donde desarrolla esta técnica estadística que es utilizada ampliamente en la construcción, análisis y validación de las pruebas.
- En 1950, Gulliksen publica un libro en el que expone y sintetiza todo lo realizado hasta entonces sobre la teoría clásica de los tests: *Theory of Mental Tests*.
- En 1954, se editan las primeras recomendaciones técnicas para el uso de las pruebas: *Technical Recomendations for Psychological Tests and Diagnostic Techniques*.

Como observamos, la historia de la psicometría es muy reciente, data de apenas hace un siglo. La primera publicación sobre las recomendaciones técnicas para el uso de las pruebas psicológicas tiene poco más de 50 años. A partir de la segunda mitad del siglo XX, la psicometría continúa su desarrollo y tiene un auge en los años ochenta que continúa hasta nuestros días. En la actualidad, existen tanto publicaciones traducidas del inglés como textos en nuestro idioma y cada día se elaboran más pruebas al mismo tiempo que se actualizan y revisan las ya existentes.

Entendemos así por psicometría el conjunto de modelos formales que establecen las bases para que se lleve a cabo de manera adecuada la medición de variables psicológicas,

con el objetivo de proporcionar los métodos para poder transformar los hechos en datos, mediante la asignación de valores numéricos tanto a las respuestas dadas por los sujetos como a los estímulos presentes en la situación de prueba, y poder interpretar esos números en función de la teoría sobre la que la prueba descansa.

Teoría clásica de las pruebas[1]

Reconocer que toda medida psicológica, al igual que las medidas de las ciencias físicas, contiene un término de error, fue un gran avance para la evaluación psicológica, ya que entonces se empieza a considerar que no basta el juicio humano para determinar los atributos psicológicos, y que los instrumentos de evaluación psicológica deben intentar, al igual que los instrumentos de las ciencias físicas, medir con precisión. Al respecto, Martínez (1996) considera que aunque en todas las ciencias el conocimiento de las propiedades de los instrumentos de medida es fundamental, el psicólogo debe poner más atención en la elaboración de sus instrumentos, ya que los atributos psicológicos no pueden medirse directamente, sino que son constructos teóricos que intentan explicar la conducta humana, por lo que el grado en que un individuo está caracterizado por esos constructos se infiere a partir de observaciones de su conducta. En este mismo sentido, enfatiza que el diseño de instrumentos para medir dichos constructos psicológicos presenta importantes problemas como:

1. No existe una sola aproximación a la medición de un constructo que sea universalmente aceptada; esto es, siempre existirá la posibilidad de que dos teóricos seleccionen diferentes tipos de conducta para la definición operativa del constructo
2. Las medidas psicológicas se basan en muestras limitadas de conducta; nunca se medirá exhaustivamente la conducta, sino sólo una muestra representativa.
3. La medida obtenida siempre tiene error; una de las principales cuestiones en psicología es estimar ese error
4. Falta de escalas con origen y unidades de medida bien definidas. No siempre una puntuación de cero significa ausencia del atributo medido; por otro lado, también pueden presentarse problemas al interpretar las medidas de acuerdo con la escala en la que están medidos sus datos
5. Los constructos psicológicos no pueden definirse aisladamente en términos de definiciones operacionales únicamente, sino que deben establecer relaciones con otros constructos y con otros fenómenos observables

La teoría de las pruebas tiene que ver con el estudio de los problemas anteriores y con la búsqueda de métodos para su solución. Mientras que la evaluación psicológica pone el acento en los contenidos sustantivos y en la interpretación de las pruebas, la base de la teoría de las pruebas se interesa por la adaptación de la estadística y el diseño experimental para dar solución a los problemas ya mencionados.

Como ya vimos, la psicometría es una rama relativamente muy joven de la psicología

que tuvo su origen en 1904, con los primeros trabajos de Spearman (Martínez, 1996; Muñiz, 2003), quien, influido por Galton y Pearson, buscaba un modelo estadístico que fundamentase las puntuaciones de las pruebas y permitiera la estimación de los errores asociados a todo proceso de medición, por lo que presenta así la primera formulación de la teoría de las puntuaciones verdaderas y el error dentro de un marco correlacional y aplicada a las puntuaciones de las pruebas; a este modelo se le conoce como Modelo Lineal Clásico de Spearman y es el que se usa en la actualidad en la teoría clásica de las pruebas. Muñiz (2003) enfatiza que no hay que perder de vista que el objetivo central de este modelo es la estimación de los errores de medición cuando se utilizan pruebas para medir variables psicológicas, y que la necesidad de un modelo tal proviene del hecho elemental de que los errores no son observables directamente. El valor que se obtiene una vez aplicada la prueba es el valor empírico mezclado con el error cuya cuantía se desea estimar.

A partir de este sencillo modelo lineal y del conjunto de supuestos en que se basa, es posible construir pruebas y analizar sus elementos, así como determinar su confiabilidad y validez; este modelo se expresa como:

X = V + e
En donde:
X es la puntuación empírica de un sujeto
V es la puntuación verdadera y
e es el término de error

La puntuación obtenida o empírica de un sujeto en una prueba (X) es igual a la puntuación verdadera (V), que nunca se puede conocer, más el error de medición (e); mientras más confiable sea la prueba y menor sea el error, la puntuación empírica se acercará más a la puntuación verdadera. Además de los errores propios del instrumento de medición, es razonable pensar que la puntuación empírica obtenida por un sujeto no coincida con su verdadera puntuación, ya que cuando se le aplica la prueba, se encuentra afectado por múltiples condiciones que no se pueden controlar y que influirán en su ejecución (fatiga, hambre, sueño, preocupación, nerviosismo, etc.). Si fuera posible lograr que el error fuera cero, entonces y sólo entonces, en ese caso la puntuación empírica sería igual a la puntuación verdadera.

Este modelo tiene tres supuestos y varias deducciones derivadas de él. Los tres supuestos son (Muñiz, 2003):

1. La puntuación verdadera de un sujeto (V) es la esperanza matemática de la empírica [E(X)]:
   V = E(X)
   La esperanza matemática sería el valor que se encontraría si se aplicara infinitas veces el mismo instrumento al mismo sujeto. Este valor sería el promedio aritmético de las puntuaciones obtenidas en las infinitas aplicaciones, si estimamos

que cada aplicación no afectara a las otras y que el sujeto no cambiara en el curso de las aplicaciones.

2. No existe correlación entre las puntuaciones verdaderas de los sujetos en una prueba y sus respectivos errores de medida:
   $\rho(v, e) = 0$
   Esto significa que los errores, como se esperaría, son aleatorios y no dependen del valor de la puntuación verdadera, son independientes de ella y por lo tanto insesgados.
3. Los errores de medida de los sujetos en una prueba no se correlacionan con sus errores de medida en otra prueba distinta:
   $\rho(e_j, e_k) = 0$
   Si las pruebas se aplican correctamente, los errores serán aleatorios y no habrá razones para que covaríen sistemáticamente unos con otros, es decir, las pruebas serán independientes.

Hay varias deducciones del modelo, aquí sólo enunciaremos cuatro de las más sencillas; si el lector quiere ver las siete restantes, puede consultar el libro de Muñiz (2003).

Deducciones del modelo:

1. El error de medida es la diferencia entre la puntuación empírica y la verdadera.
   El modelo lineal establece que:
   $X = V + e$
   Si despejamos **e,** tenemos que:
   $\mathbf{e = X - V}$
2. La esperanza matemática de los errores de medida es cero, por lo tanto, son errores insesgados.
   De la deducción anterior, tenemos que:
   $e = X - V$
   La esperanza matemática del error entonces es:
   $E(e) = E(X) - E(V)$
   Pero como sabemos, por el supuesto 1, que
   $V = E(X)$
   Si sustituimos en la ecuación anterior, tenemos que:
   $\mathbf{E(e) = V - E(V) = V - V = 0}$
3. La media de las puntuaciones empíricas es igual a la media de las puntuaciones verdaderas.
   De acuerdo con el modelo de que $X = V + e$:
   $E(X) = E(V + e) = E(V) + E(e)$
   Pero como ya vimos $E(e) = 0$, si sustituimos:
   $E(X) = E(V) + 0 = E(V)$
   Por lo tanto, si $E(X) = E(V)$, entonces:

$\mu_x = \mu_v$

4. Las puntuaciones verdaderas no covarían con los errores, por lo tanto, los errores son insesgados.
   Por definición, la covariación entre V y e es:
   $cov\ (V,e) = \rho_{ve}\sigma_v\sigma_e$
   Según el supuesto 2 del modelo:
   $\rho(v,\ e) = 0$
   Sustituyendo:
   $\mathbf{cov\ (V,e) = (0)\ \sigma_v\sigma_e = 0}$

Los supuestos y las deducciones anteriores significan que, de acuerdo con el modelo lineal de la teoría clásica de las pruebas, al aplicar un instrumento elaborado bajo este marco esperaríamos teóricamente que la puntuación obtenida por un sujeto fuera igual a la puntuación verdadera, con un margen de error aleatorio que tendería a ser cero y que no tendría ninguna relación con la puntuación obtenida en la prueba, ni con algún otro instrumento aplicado al mismo sujeto.

A nivel empírico, si el instrumento está bien construido, esperaríamos que la puntuación obtenida por el sujeto refleje lo mejor posible su puntuación verdadera y el error de medición sea cercano a cero; de acuerdo con la teoría clásica de las pruebas, esperamos que en la medición de atributos psicológicos, la puntuación obtenida por un sujeto en una prueba se acerque lo más posible a su ejecución real con el mínimo posible de errores, esto es, que el instrumento sea confiable para medir el atributo que nos interesa. Así, para poder considerar a un instrumento de evaluación psicológica como adecuado y científico, debe contar, como los instrumentos de todas las ciencias, con dos requisitos indispensables: la confiabilidad y la validez.

Confiabilidad

La confiabilidad no es un asunto de todo o nada, sino una cuestión de grado: es un continuo que abarca desde la consistencia mínima de una medición a la casi perfecta repetibilidad de los resultados; las pruebas psicológicas se encuentran en algún lugar dentro de este continuo (Muñiz, 2003). Una medida es confiable si conduce a los mismos resultados o a resultados muy similares, "…sin importar las oportunidades para que ocurran variaciones" (Nunnally y Bernstein, 1995, pág. 238), es decir, la confiabilidad se refiere a la capacidad que tiene una medida obtenida, de poder generalizarse a otras situaciones.

En un sentido amplio, la confiabilidad es la exactitud, la precisión con que un instrumento mide un objeto; en términos estrictos, la confiabilidad sería la ausencia de errores de medición; sin embargo, en psicología, es difícil utilizar este término en la medición de atributos psicológicos, como alternativa se utilizan sinónimos como estabilidad de la medida y consistencia interna, que explicaremos más adelante.

Como ya veíamos en el modelo lineal clásico de la teoría de las pruebas ($X = V + e$), el

término **e** o error de medición es todo aquello que, como parte de la puntuación obtenida por el sujeto en la prueba, difiere de la puntuación verdadera; la confiabilidad será más alta mientras más bajo sea ese término de error, ya que entonces la puntuación obtenida por el sujeto en una prueba se acercará más a su puntuación verdadera. En otras palabras, la puntuación obtenida por una persona en una prueba es una mezcla de la cantidad del atributo que posee dicha persona y el error de medición. Pero, ¿qué es lo que hace que el error aumente o disminuya en una medición?, ¿cuáles son las fuentes de error más comunes cuando se evalúa a una persona?

Fuentes de error en la medición en psicología

La primera fuente de error en la evaluación psicológica es la **selección de los reactivos**. El autor de una prueba, una vez que elabora un conjunto de reactivos para medir el atributo meta, debe decidirse por seleccionar una muestra de ellos; ¿cómo elegirlos para que realmente sea representativa, para que sea equitativa en cuanto a los diferentes rubros que su prueba tiene? Por ejemplo, si una prueba de aritmética tiene más reactivos de sumas que de restas y un alumno estuvo ejercitándose más en sumas que en restas, su calificación será más alta que la que hubiera obtenido si se hubiera ejercitado más en restas que en sumas, o bien, si la prueba tuviera igual número de reactivos de ambas operaciones aritméticas. Una prueba bien diseñada y construida, debe asegurar que la muestra de reactivos contenga todos los tipos de contenidos que intenta medir, y de manera equitativa.

Otra fuente de error es la **aplicación de la prueba**. Los manuales de las pruebas presentan las condiciones estandarizadas de aplicación de la misma, así como las instrucciones que se les deben dar a los sujetos. Sin embargo, a veces hay factores que no se pueden controlar: un espacio no muy bien iluminado, libre de ruidos y distractores, mesa y silla cómodas, etc., así como las condiciones físicas del que está tomando la prueba: cansancio, falta de motivación, ansiedad, preocupación, hambre, sueño, sed y, en general, los estados emocionales; no siempre es posible controlar estos factores y forman parte del error aleatorio de medida.

La **calificación de la prueba** constituye otra fuente de error. Si bien existen claves o plantillas de corrección de las pruebas, en algunas de ellas hay reactivos cuya respuesta tiene que valorar y calificar el evaluador; por ejemplo, en las pruebas de inteligencia, cuando existen preguntas abiertas como: ¿qué harías si te encontraras una mochila tirada en el patio de tu escuela?, el evaluador tiene que decidir, de acuerdo con la respuesta dada por el evaluado, si le otorga la calificación de 0, 1 ó 2 puntos. Los constructores de las pruebas reducen este tipo de error al proporcionar en los manuales de las mismas muchos ejemplos de los posibles tipos de respuesta para que cualquier evaluador tenga más certeza en otorgar alguna puntuación.

A los tipos de errores que hemos mencionado se les conoce como **errores aleatorios o errores no sistemáticos,** porque sus efectos son inconsistentes e imposibles de predecir; como su nombre lo indica, contribuyen de manera aleatoria a la puntuación de los

sujetos, a veces lo favorecen y otras lo perjudican, por lo cual, ya que son aleatorios, al sumarse, tenderían a acercarse a cero (puntuaciones negativas más positivas) y su efecto sobre la confiabilidad sería no significativo. Sin embargo, existen otros errores, los sistemáticos, que son los que realmente le preocupan a los constructores de pruebas, ya que atentan de manera significativa contra la confiabilidad del instrumento. Un error sistemático surge cuando, por ejemplo, la prueba mide de manera consistente algo diferente del propósito de la prueba (Gregory, 2001), como cuando un reactivo está mal planteado, es poco claro o induce la respuesta. El error sistemático a su vez puede ser *constante*, esto es, afectar a todas las puntuaciones por igual, o puede ser un *sesgo*, y afectar a ciertos tipos de observaciones de manera diferente que a otras; Nunnally y Bernstein (1995) nos dan un buen ejemplo de error sistemático constante en un instrumento de medición de la medicina: un termómetro mal calibrado que siempre marcara tres grados de más.

La manera de minimizar los errores sistemáticos y, por lo tanto, maximizar la confiabilidad es seguir las recomendaciones técnicas de los psicómetras para la elaboración de pruebas; no obstante, por mucho cuidado que se tenga en la elaboración de los reactivos, el único modo de asegurar que la prueba puede ser utilizada de una manera confiable es, una vez construida, obtener mediante técnicas estadísticas su coeficiente de confiabilidad. En este punto hay que agregar que para que una prueba se publique, los editores exigen datos sobre su confiabilidad y validez, de esta manera, podemos estar seguros de que las pruebas que existen en el mercado son confiables y las podemos aplicar a los sujetos que así lo requieran, no así pruebas que se bajan de internet o se consiguen en fotocopias.

Confiabilidad como estabilidad de la medida

Si midiéramos un objeto físico, por ejemplo, una mesa con un instrumento *ad hoc* para medir longitudes, como una cinta métrica graduada en centímetros y con divisiones para milímetros, el error de medición que podríamos cometer —suponiendo que elimináramos al máximo los errores aleatorios como colocar la cinta métrica inadecuadamente, ver la medida de lado y no de frente, mover la cinta a la hora de registrar la medida, etc.— es del orden de 0.1 a 0.4 cm a favor o en contra: si el final de la mesa cae después de la raya que indica 50 cm, pero antes de 50.5 cm, diríamos que esta medida podría ser 50.1, 50.2, 50.3 o 50.4 cm, según nuestra apreciación. ¿Cuál sería la mejor medida? Si en lugar de hacer una sola medición hacemos varias —digamos 10, ya sea por una sola persona, o bien, por 10 personas diferentes—, y si el instrumento utilizado es confiable, observaríamos que estas mediciones son muy cercanas, por ejemplo: 50.3, 50.4, 50.4, 50.2, 50.3, 50.3, 50.3, 50.2, 50.3, 50.3, esto es, la medida es estable a través de las diferentes mediciones, con un margen de variación muy pequeño (de 0.2 a 0.4 cm); si las mediciones difirieran mucho unas con otras, entonces el instrumento no sería confiable al no haber consistencia entre las mediciones.

La estabilidad de la medida se refiere a que al medir un atributo psicológico con un

determinado instrumento, éste será confiable si al evaluar a los mismos sujetos con el mismo instrumento o con uno equivalente, las medidas obtenidas en la segunda aplicación son muy similares a las obtenidas en la primera, esto es, son estables a través del tiempo, lo cual indicaría que los errores de medición serían mínimos y, por lo tanto, la confiabilidad sería aceptable; las diferencias encontradas entre una medición y otra se atribuirían a los errores aleatorios asociados al proceso de medición y no al instrumento (Muñiz, 2003). Sin embargo, cuando medimos constructos psicológicos, no esperamos un grado de error tan pequeño como cuando medimos objetos físicos.

Ya que en psicología no tenemos instrumentos como los de las ciencias físicas que miden con tanta precisión, uno de los modos de obtener la confiabilidad de los instrumentos psicológicos es mediante la estabilidad de la medida: si aplicamos un instrumento a una muestra determinada de personas y lo volvemos a aplicar después de un tiempo para obtener la confiabilidad, las medidas obtenidas entre la primera y la segunda aplicación no deben ser tan diferentes, para poder acreditar al instrumento como confiable. Así, para conseguir este tipo de confiabilidad existen dos técnicas: obtención de la confiabilidad por los métodos test-retest y formas equivalentes de prueba o pruebas paralelas. Sin embargo, cuando utilizamos un instrumento para medir un objeto físico repetidas veces, como en el ejemplo anterior de la mesa, el objeto puede medirse en varias ocasiones sin problemas, ya que la mesa no cambia de longitud con el paso del tiempo, a diferencia de los atributos psicológicos, donde puede haber inconsistencias atribuidas tanto a la falta de confiabilidad del instrumento, como a cambios inevitables que pudieran ocurrir en el sujeto evaluado o en la situación de evaluación.

En la confiabilidad test-retest se aplica la prueba a una muestra y se vuelve a aplicar un tiempo después, que puede variar entre una semana y varios meses. Teóricamente, si la prueba careciera totalmente de errores de medición y el paso del tiempo no fuera un factor importante en el atributo a medir, se esperaría que cada sujeto evaluado obtuviera exactamente la misma puntuación. Si éste fuera el caso, al graficar los datos y obtener el coeficiente de correlación r de Pearson, tendríamos una gráfica como la siguiente:

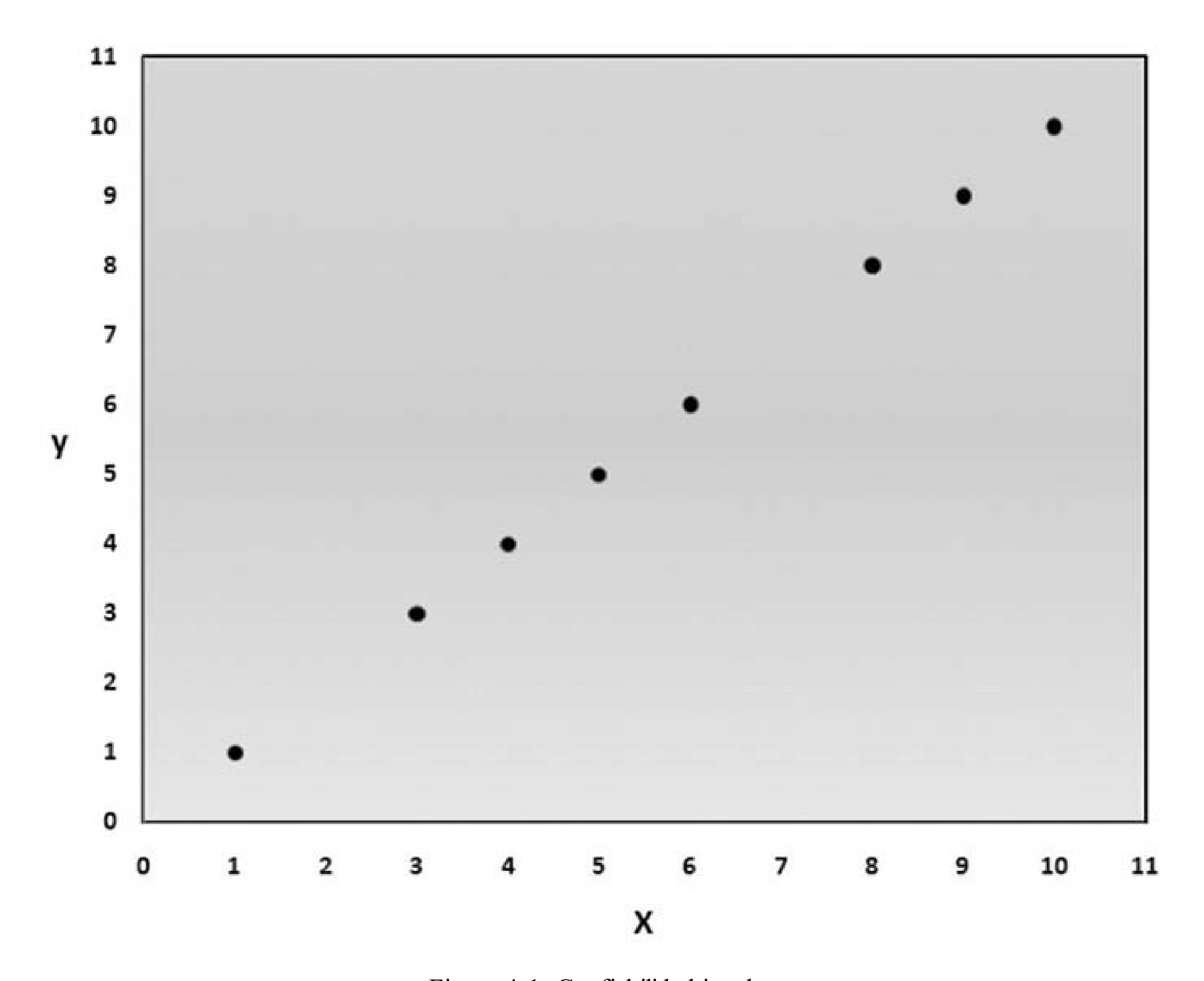
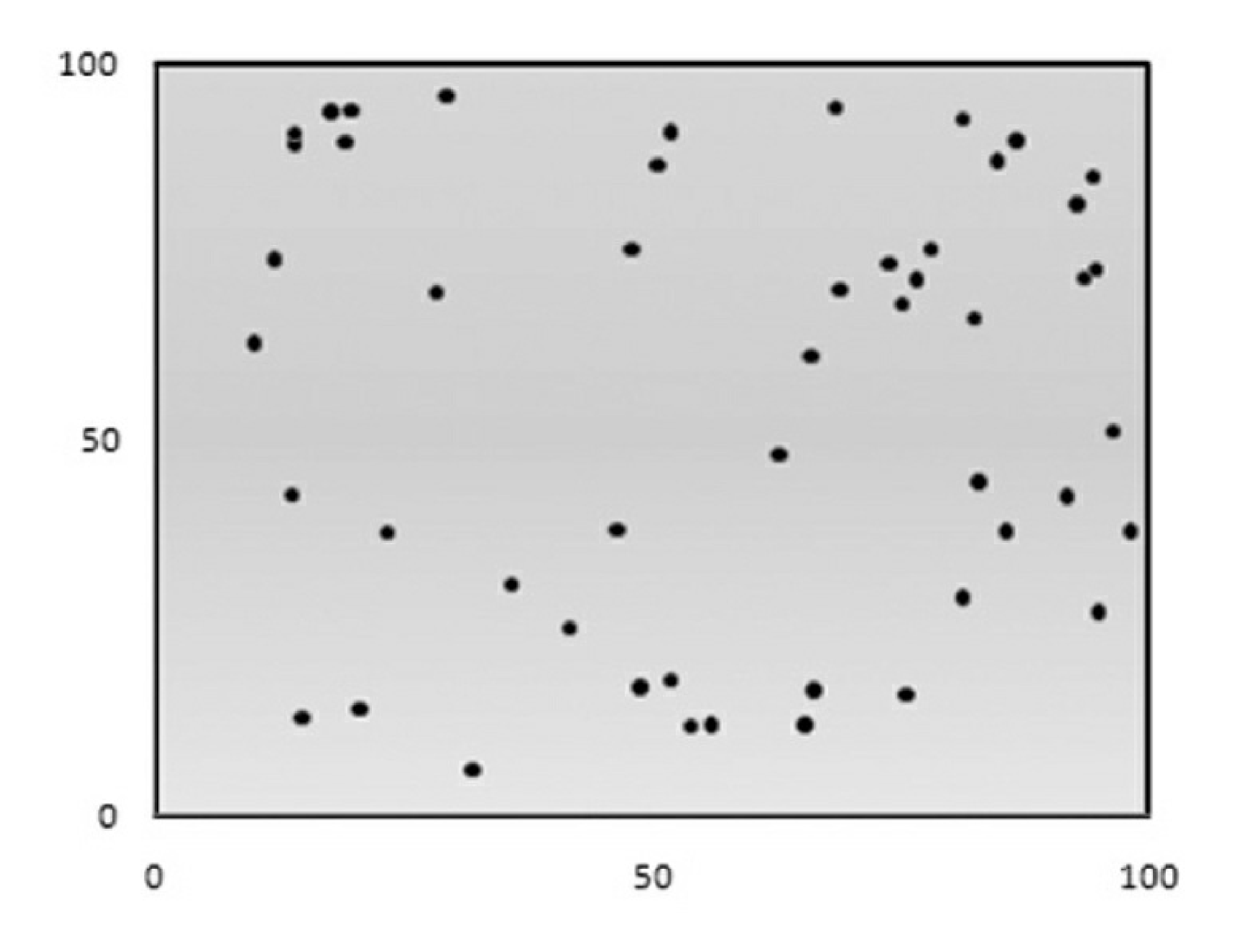


Figura 4-1. Confiabilidad igual a uno

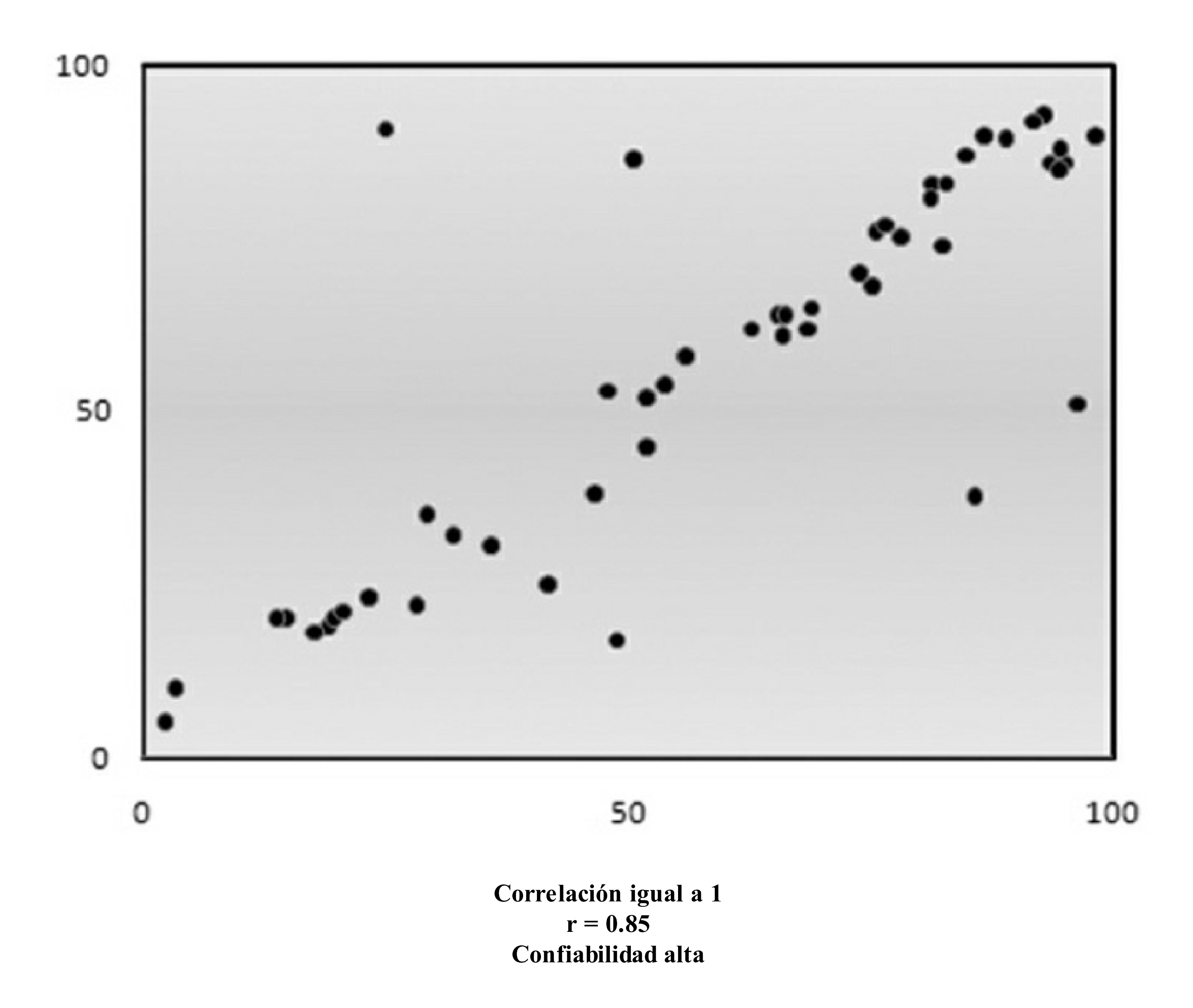
En esta Figura 4-1, al puntaje de cada sujeto en la primera aplicación (eje X) le corresponde exactamente la misma puntuación en la segunda (eje Y). Al obtener el valor de la correlación entre ambos pares de puntuaciones, alcanzaríamos un valor r=1, que sería la correlación perfecta.

De esta manera, el índice de confiabilidad de un instrumento de evaluación medido como estabilidad temporal, se calcula por medio del coeficiente r de Pearson, cuyos valores oscilan entre 0 —que indicaría total ausencia de confiabilidad— hasta 1, que sería la confiabilidad perfecta, la cual no existe, pero mientras más se acerque el valor de la r de Pearson a 1, más confiable es el instrumento.



Correlación igual a cero
r = 0
Nula confiabilidad

Figura 4-2. Confiabilidad igual a cero



Correlación igual a 1
r = 0.85
Confiabilidad alta

Figura 4-3. Confiabilidad igual a 0.85

El valor de la confiabilidad indica el porcentaje de varianza en las puntuaciones obtenidas, que es explicado por la variabilidad en las puntuaciones verdaderas y en qué medida se explica por efectos aleatorios. Por ejemplo, un valor de confiabilidad de 0.85 como el de la Figura 4-3, indicaría que la puntuación del sujeto se explica en 85% por las puntuaciones verdaderas obtenidas y en 15% por causas no determinadas, es decir, el instrumento es confiable en 85%. Por otro lado, la Figura 4-2 nos ilustra un caso de total ausencia de confiabilidad (r=0), no hay en absoluto concordancia entre las puntuaciones obtenidas en la primera y la segunda aplicación de la prueba.

Sin embargo, obtener la confiabilidad de esta forma tiene algunas desventajas; por ejemplo, si se trata de una prueba de inteligencia y entre las dos aplicaciones hay una diferencia de tiempo corta (una semana), el efecto de memoria contribuirá a que los sujetos evaluados contesten de igual forma que la anterior, además de que la primera aplicación les pudo haber servido de práctica y entonces obtener mejores puntajes; por otro lado, si se deja transcurrir mucho tiempo, los inconvenientes serían los efectos de

maduración y aprendizaje, que podrían elevar las puntuaciones obtenidas en la primera aplicación. Para evitar estas amenazas que compiten con la medición de la confiabilidad, se podría disponer de dos pruebas equivalentes o paralelas; sin embargo, al elaborarlas nunca tendríamos la certeza de que realmente posean el mismo grado de dificultad para medir el atributo en cuestión, además del alto costo y tiempo que implicaría elaborar dos pruebas psicológicas. Por ello, los constructores de pruebas le siguen dando prioridad a la técnica *test-retest* para obtener la confiabilidad de sus instrumentos de medición.

Confiabilidad como consistencia interna

Por otro lado, la consistencia interna se refiere a que los reactivos de un instrumento dado son consistentes entre sí en la forma en que evalúan el atributo psicológico propuesto: los sujetos que poseen el atributo medido por la prueba en cantidad elevada, tendrán un puntaje alto en los reactivos de la misma, y los sujetos que no poseen el atributo o lo tienen en poca cantidad, tendrán un puntaje bajo. Esta forma de obtener la confiabilidad tiene la ventaja de que no es necesario aplicar la prueba dos veces, sino que con una basta.

Hay dos formas de obtener la confiabilidad desde esta perspectiva: confiabilidad de división por mitades y consistencia interna de los reactivos individuales. En la primera, se correlacionan los reactivos de una mitad de la prueba -generalmente los reactivos nones- con los de la otra mitad -los reactivos pares-, utilizando para tal fin la fórmula de correlación corregida Spearman-Brown, o bien, si la puntuación de los reactivos de la prueba es dicotómica, la fórmula Kuder-Richardson; se parte del supuesto de que ambas mitades son equivalentes y es como si se aplicaran dos pruebas cortas equivalentes. Este tipo de confiabilidad se utiliza en pruebas que miden habilidades intelectuales. Supongamos que los reactivos se enlistan en la prueba en orden de dificultad creciente y se seleccionan así para correlacionar los reactivos pares con los reactivos nones, si la correlación obtenida es alta, se infiere una alta confiabilidad en la prueba; sin embargo, la crítica a la obtención de la confiabilidad por este método recae en la supuesta equivalencia de ambas mitades de la prueba, además de que es dudosa su utilización en pruebas que miden constructos del área emocional o no intelectual. Cronbach (1951, citado en Gregory, 2001) acotó que para no depender de una sola división, se podría obtener la media de los coeficientes por mitades de todas las posibles formas de dividir la prueba en dos, lo que dio paso a la confiabilidad medida por el coeficiente Alfa de Cronbach.

En la segunda forma de obtener la confiabilidad, se mide la consistencia interna de los reactivos individuales por medio del coeficiente Alfa de Cronbach, que se considera como la media de todas las correlaciones que pudieran obtenerse al dividir la prueba en todas las posibles mitades. Para que quede más claro el concepto de consistencia interna, supongamos que estamos elaborando una prueba para medir hábitos de estudio, con dos versiones, en la primera versión se incluye el reactivo 8a y en la segunda el reactivo 8b. En esta prueba, se otorga un punto por cada reactivo par contestado afirmativamente y

por cada reactivo non contestado negativamente. Los reactivos son los siguientes:

1. ¿El lugar donde estudia es incómodo y con mucho ruido?
2. ¿Deja sus problemas personales cuando va a estudiar?
3. ¿A la hora de realizar un trabajo se da cuenta de que no tiene todo a la mano?
4. ¿Subraya las ideas más importantes del texto?
5. ¿Le da pena preguntar al profesor cuando no entiende algo en clase?
6. ¿Busca en el diccionario las palabras que no entiende cuando está leyendo un texto?
7. ¿Se aprende de memoria lo que no entiende?
8a. ¿Acostumbra hacer resúmenes o esquemas cuando está estudiando?
8b. ¿Sus zapatos actuales son incómodos?

En este ejemplo, en la primera versión, si el sujeto contesta afirmativamente los reactivos 2, 4, 6 y 8a, y negativamente los reactivos 1, 3, 5 y 7, obtendrá la puntuación más alta; a más alta puntuación, mejores hábitos de estudio. Aquellos sujetos que tienen puntaje alto en la prueba tenderán a responder afirmativamente a los reactivos 2, 4, 6 y 8a, y negativamente a los reactivos 1, 3, 5 y 7; mientras que los que tienen puntaje bajo en general en toda la prueba tenderán a responder afirmativamente a los reactivos 1, 3, 5 y 7 y negativamente a los reactivos 2, 4, 6 y 8a. Al dividir la prueba en todas las mitades posibles en este caso, el número de todas las formas o combinaciones posibles de ocho elementos o reactivos tomados de cuatro en cuatro es igual a 70[2], por lo que los pares de mitades para correlacionar serían 35, y obtener la media de las 35 parejas de correlaciones, como los reactivos miden consistentemente el atributo medido: hábitos de estudios, seguramente se obtendría un valor alto del coeficiente de correlación Alfa de Cronbach. Sin embargo, si aplicáramos la segunda versión de la prueba, con el reactivo 8b, a una población determinada y correlacionáramos las puntuaciones obtenidas de las 35 posibles parejas de reactivos, aquella mitad donde quede ubicado el reactivo 8b (*¿sus zapatos actuales son incómodos?*), no correlacionará positivamente alto con la otra mitad, debido a que el reactivo 8b pudiera ser contestado afirmativa o negativamente, tanto por un alumno con buenos hábitos de estudio como por aquél que no los tiene, esto es, ese reactivo es inconsistente, comparado con los demás reactivos, para medir el atributo de interés. El coeficiente Alfa de Cronbach nos proporcionaría, además de un índice de correlación, el conocimiento de qué reactivo es el que se está comportando de manera inconsistente y por lo tanto produce un valor más bajo de confiabilidad que el que se esperaría si se eliminara de la prueba.

—Aclaramos que éste no es un libro de estadística —por eso no proporcionamos las fórmulas para encontrar los diferentes coeficientes de confiabilidad— sino para que el lector comprenda, cuando tenga en sus manos el manual de una prueba y revise los datos sobre confiabilidad, el tipo de coeficiente utilizado para conseguirla y cómo la obtuvo el constructor, además de interpretar el resultado final de dicho coeficiente.

Así, podemos resumir que una prueba es confiable en sentido psicométrico si

consistentemente produce, al ser aplicada en repetidas ocasiones, la misma puntuación o una muy similar, o bien, si los reactivos que la componen son consistentes entre sí en la forma en que miden el atributo propuesto por la prueba; veamos ahora el significado de esa medida.

Validez

En un sentido muy general, un instrumento de medición es válido si hace aquello para lo que fue concebido; la validez de una prueba concierne a lo que ésta mide, su eficacia y lo que podemos inferir de los puntajes obtenidos en la prueba. Una medida puede ser confiable sin ser válida, por ejemplo, podemos utilizar una cinta métrica y medir en una jarra el nivel en donde se encuentra el agua (p. ej. 10.5 cm); esta medida es confiable, pero no válida: si cambiamos el agua a otro recipiente la medida será diferente y no podremos generalizar lo medido, no es válido medir capacidad con un instrumento que mida longitud. Sin embargo, para que una medida sea válida, necesita primero ser confiable; no tendría sentido hablar de validez de un instrumento de medición si no mide el atributo de manera confiable. La confiabilidad es una condición necesaria pero no suficiente para la validez.

El estudio de la validez de un instrumento se refiere básicamente a validar los datos proporcionados por éste: el grado de adecuación, significación y utilidad de las inferencias específicas que pueden derivarse a partir de las puntuaciones de las pruebas; como menciona Martínez (1996), debemos tener siempre presente que lo que se valida no es el instrumento, sino la interpretación de los datos obtenidos por medio de un procedimiento específico, las inferencias que podemos deducir de la ejecución de un sujeto en una prueba, es decir, si la conducta mostrada en la situación de prueba es o no un reflejo de la conducta habitual del sujeto en situaciones naturales.

La validación requiere siempre de investigaciones empíricas y el tipo de datos necesarios para ello depende de la clase de validez que se desea establecer y del uso que se les dará a los datos obtenidos con la aplicación de la prueba; al igual que la confiabilidad, la validez no es un asunto de todo o nada, sino una cuestión de grado. Ninguna prueba es válida en general, una prueba sólo es válida para la finalidad específica para la que fue construida, por lo tanto, puesto que las pruebas se emplean para diversos fines, no existe un solo tipo de validez para todos los propósitos de evaluación y en el evaluador recae la responsabilidad del empleo válido de los resultados de la prueba. Fundamentalmente, todos los procedimientos para determinar la validez de una prueba conciernen a las relaciones entre la ejecución y otros hechos observables de manera independiente acerca de las características de la conducta que se estudia. Los tipos de validez que los *Standards of the American Psychological Association* han establecido, según el tipo de prueba y el uso al que esté destinada, son (Anastasi, 1988): validez de constructo, validez de criterio y validez de contenido; a continuación se describen cada una de ellas.

Validez de constructo

En la medida en que alguna variable es abstracta, hablamos de ella como un constructo. Una variable así es literalmente un constructo, pues es algo que no existe como dimensión de conducta observable; cada constructo se desarrolla para explicar y organizar consistencias de respuestas observadas (p. ej. inteligencia o personalidad). La validez de constructo es la extensión en la cual la prueba dice medir un constructo o rasgo teórico y requiere la acumulación gradual de información de diferentes fuentes; cualquier dato que arroje luz sobre la naturaleza del constructo bajo consideración y de las condiciones que afectan su desarrollo y sus manifestaciones, representa una evidencia apropiada para este tipo de validación (Martínez, 1996). Ejemplos de pruebas que miden constructos son las de inteligencia, personalidad y las que miden la ansiedad como rasgo (como la prueba IDARE).

La validez de constructo se basa en el significado psicológico del puntaje de una prueba y en la explicación teórica de la ejecución del sujeto. Cuando se indaga la validez de los constructos de las pruebas psicológicas, se intenta saber qué propiedades psicológicas o de otra índole pueden explicar la varianza de esas pruebas, es decir, explicar las diferencias individuales observadas en las puntuaciones del instrumento; así, no se trata sólo de validar la prueba, sino que se valida también la teoría sobre la cual ésta descansa (Kirsch y Guthrie, 1980). Puesto que la medición de constructos es una parte vital de la evaluación psicológica, ¿cómo se establecen y validan esas medidas? En el proceso, hay tres aspectos principales: 1) especificación del dominio de las conductas observables, 2) determinar hasta qué punto todas o algunas de esas conductas se correlacionan entre sí y 3) precisar si una, algunas o todas las medidas de tales conductas actúan, como si midieran el constructo (Nunnally, 1973).

La prueba suficiente de la validez de constructo es que las medidas del constructo se comporten como se espera. Por ejemplo, si se supone que una prueba determinada mide el constructo "ansiedad", el sentido común sugiere que los puntajes mayores en esta prueba, que indicarían mayor ansiedad, se encontrarían en el caso de pacientes clasificados como "neuróticos ansiosos" con mayor probabilidad que en pacientes no clasificados como tales; en sujetos de un experimento a quienes se les amenaza con un *shock* eléctrico y no en el caso de aquellos a quienes no se les amenaza; en estudiantes de licenciatura antes de tomar un examen final oral de alguna materia más que a esos mismos estudiantes una vez que han pasado su examen. Otro ejemplo: si se establece que una determinada prueba mide el constructo "inteligencia", se esperaría que correlacionara por lo menos moderadamente con las calificaciones escolares, las consideraciones acerca de la inteligencia hechas por profesores y con los niveles de realización profesional. Así sucede con todos los constructos: se esperan ciertas relaciones con otras variables. Para obtener un índice numérico de la validez de constructo se utiliza también la r de Pearson, que relaciona los valores obtenidos en la prueba con medidas que se supone teóricamente se correlacionan con el constructo (como la inteligencia y las calificaciones escolares), o bien, con otra prueba que mida el

mismo constructo y que ya esté validada.

En resumen, el propósito de la validez de constructo es el de validar la teoría subyacente al sistema de evaluación y a la medida misma. Los constructos son los que determinan qué conductas han de seleccionarse para su observación y la situación en la que ésta será aplicada. Un instrumento de medida estará ligado al sistema de constructos dentro del que fue construido; este tipo de validez nos indicará el grado en que el instrumento de evaluación es una medida adecuada del constructo y en qué grado o alcance también las hipótesis derivadas del mismo pueden confirmarse mediante la utilización del instrumento en cuestión.

Validez referida al criterio

La validez referida al criterio o predictiva es aquella que se utiliza para estimar a futuro una conducta, a la que se llama criterio; una vez determinado el criterio, se elaboran los reactivos que estarán correlacionados con él. Para obtener el valor de la validez del criterio, se correlacionan los puntajes de la prueba con los puntajes de la variable criterio.

La validez referida al criterio valora el grado en el que un instrumento de evaluación puede utilizarse para estimar la conducta de una persona en una situación concurrente con la aplicación de la prueba (validez concurrente o diagnóstica), como asignar un diagnóstico clínico, tal como esquizofrenia, hipocondriasis, trastorno obsesivo-compulsivo[3], o bien, estimar su conducta en situaciones futuras (validez predictiva), como pudieran ser sus aptitudes para el aprendizaje, preferencias vocacionales, asignación o selección de personal. Así, se compara la ejecución de los sujetos en la prueba con un criterio, que es una medida directa e independiente de lo que la prueba predice (Martínez, 1996); se correlacionan los puntajes obtenidos en una prueba (p. ej. la Batería de aptitudes para el aprendizaje escolar, que se aplica al finalizar la educación preescolar y que predice si el alumno tendrá éxito para aprender a leer y escribir, y para la aritmética básica al finalizar el primer año de primaria), con las puntuaciones a futuro que se obtendrán en otra prueba o con un hecho a futuro (p. ej., los exámenes finales del curso de primer año o las calificaciones finales). En este tipo de validez, lo más importante es la capacidad predictiva de la prueba.

En resumen, la validez referida al criterio se caracteriza por la predicción relacionada con un criterio externo y porque recurre a la comprobación del instrumento de medición, ya sea en el momento presente o en el futuro, y lo compara con un resultado o medida.

Validez de contenido

Para algunos instrumentos, la validez depende primordialmente de la suficiencia con que se muestrea un dominio específico de contenido; la prueba debe ser una medida adecuada de lo que evalúa. La validez de contenido es la representatividad o adecuación muestral del contenido del instrumento de medición, es decir, es una clara descripción del dominio de conductas de interés (Linehan, 1980).

Toda propiedad psicológica posee un universo teórico de contenido constituido por todo aquello que se puede afirmar u observar acerca de ella. La validez de contenido responde a la pregunta: ¿la sustancia o contenido de este instrumento de medición es representativa del contenido o del universo de contenido de la propiedad que se va a medir? (Martínez, 1996).

La validez de contenido incluye esencialmente el examen sistemático del contenido de la prueba para determinar si cubre una muestra representativa del dominio de conductas que se medirán; lo que interesa en este tipo de validez es si las respuestas dadas a los reactivos de estímulo y las condiciones bajo las cuales la conducta es observada representan a todos aquellos conjuntos de estímulos, respuestas y condiciones, a los cuales se interesa generalizar. La validez de contenido tiene como objetivo demostrar que los reactivos de la prueba son una muestra representativa de un universo y asegurar un muestreo cuidadoso de un dominio de contenido relevante. Al respecto, Martínez (1996) menciona que existen dos aspectos importantes y complementarios de la validez de contenido que deben ser tomados en cuenta en la construcción de instrumentos: 1) que el instrumento no incluya aspectos irrelevantes de la conducta de interés y 2) que el instrumento contenga todos los aspectos importantes que definen el dominio conductual.

La validez de contenido es el grado en que un conjunto de reactivos representa adecuadamente un dominio o universo de conductas. El contenido de la prueba es una muestra representativa de la clase de situaciones o problemas sobre los que se extraerán las conclusiones, lo cual garantiza que los resultados del sujeto en la prueba se pueden generalizar al universo de contenido que la prueba representa, por ejemplo, rendimiento escolar, hábitos de estudio, errores de tipo disléxico, ansiedad, habilidades sociales. Ejemplos de pruebas con validez de contenido son: la prueba de percepción visoespacial de Frostig DTVP-2, la prueba de ansiedad para niños CMASR-2, la que mide errores de tipo disléxico IDETID-LEA.

Hoste (1981) menciona que no existe un modo aceptable para cuantificar la validez de contenido y, por lo tanto, no puede ser expresada como un coeficiente de validez, por lo que se determina por medio de procedimientos racionales en vez de empíricos, con un examen cuidadoso de los procedimientos de construcción de la prueba. Por lo general, se utiliza la técnica de "jueceo", en donde jueces competentes en el atributo que se pretende medir y en el objetivo de la prueba, ponderan el contenido de los reactivos (Kerlinger, 1975).

Al igual que en el rubro de confiabilidad, es preciso aclarar que este texto no pretende que el lector elabore una prueba y sepa cómo obtener la confiabilidad y la validez, sino que, una vez que seleccione una prueba para aplicarla a algún paciente, se cerciore de que tiene datos de confiabilidad y validez, y conozca qué tipo de inferencias podrá obtener de los resultados. A continuación esbozaremos de manera breve un intento de clasificación de las pruebas psicológicas.

Clasificación de las pruebas psicológicas

No todos los instrumentos de evaluación psicológica son iguales; podemos clasificarlos según su interpretación, forma de comparación, finalidad, modo de aplicación, forma de aplicación y área del comportamiento que miden.

De acuerdo con la **interpretación,** las pruebas se dividen en objetivas y subjetivas. Las **pruebas objetivas** nos proporcionan tablas y medidas de comparación de la ejecución de los sujetos y nos permiten interpretar los resultados de una manera clara, objetiva, sin ambigüedades y sin depender del juicio del evaluador. Así, los resultados de la ejecución de un sujeto en una prueba serían interpretados de la misma manera por dos evaluadores diferentes, la interpretación de su ejecución sería igual, independientemente del evaluador (p. ej., Prueba de Inteligencia WISC-IV, Prueba de Personalidad 16PF, Prueba MMPI, Prueba de percepción visoespacial DTVP-2 de Frostig).

Por otro lado, las pruebas **subjetivas o proyectivas** son aquellas en las que la interpretación de la ejecución del evaluado descansa en los juicios del evaluador, que a su vez dependen de la teoría sobre la que la prueba se sustenta. Aunque hay líneas generales de interpretación, los resultados de un mismo sujeto podrían ser interpretados de diferente manera por distintos evaluadores (p. ej., Prueba de las Manchas de Tinta de Rorschach, Prueba del Dibujo de la Figura Humana, Prueba de la Casa, el Árbol y la Persona). Este tipo de pruebas no se construyen de acuerdo con los principios psicométricos ya revisados, por lo que no muestran datos ni de confiabilidad ni de validez, y la mayoría de ellas no convierte las respuestas de los sujetos en datos o números, es decir, no mide, sólo interpreta de manera cualitativa las respuestas de los evaluados (por lo general, en el área clínica).

Las pruebas objetivas, por la **forma de comparación,** pueden ser referidas a la norma o al criterio. En las **pruebas referidas a la norma,** la ejecución de un sujeto se compara con la ejecución de un grupo o de una población, es decir, con un grupo normativo. Por ejemplo, si al aplicar una prueba de inteligencia a un niño de siete años, éste obtiene un CI de 85, un percentil de 25 y una edad escalar de 6:4, significa que, comparado con la población de niños de siete años, su CI se encuentra por debajo de la norma esperada para su edad —que debería estar entre 90 y 110—, que su ejecución en esta prueba se encontró por abajo de 75% de los niños de su edad y que su ejecución en este atributo es comparable a la de un niño de seis años cuatro meses, esto es, ocho meses por debajo de su edad cronológica.

En las **pruebas referidas al criterio,** la ejecución de cada individuo se compara con un criterio no normativo, arbitrario, determinado de antemano, para observar qué tanto se aleja o se acerca del criterio establecido, criterio que si bien es arbitrario en el sentido de que no es normativo, se basa en la experiencia y en cuestiones teóricas. Por ejemplo, un profesor puede determinar que para acreditar su materia los alumnos deben poseer un mínimo de conocimientos que equivale a siete de calificación; al trabajar con un niño con síndrome de Down, se puede establecer como criterio para pasar a un siguiente objetivo de lenguaje que debe obtener 75% de respuestas correctas en una prueba de lenguaje; al evaluar a un niño con el IDETID-LEA (instrumento para detectar errores de tipo disléxico), se determinaría qué tipo de errores comete en copia, lectura y dictado, y en

qué universos de generalización se presentan. Así, en la evaluación de estos casos, se determinaría si alcanzan el criterio señalado, y si no, qué tan lejos se encuentran de él para iniciar o seguir trabajando en la consecución del objetivo.

Por su **finalidad,** las pruebas pueden ser utilizadas para la detección, investigación, diagnóstico, selección, predicción, planeación de un tratamiento, intervención, valoración del cambio, etc. Independientemente del objetivo de su aplicación, la finalidad última de toda aplicación de pruebas es la toma de decisiones. El objetivo perseguido en la aplicación de una prueba lo establece el evaluador, de acuerdo con sus intereses o con los de los clientes que acuden a solicitar sus servicios.

Por su **modo de aplicación,** las pruebas pueden ser de lápiz y papel, orales, de manipulación o ejecución, mixtas y computarizadas. Por su **forma de aplicación,** pueden ser individuales, colectivas y mixtas. Los manuales de las pruebas explican las instrucciones de aplicación, las cuales, a menos que se indique otra cosa, deben seguirse rigurosamente al pie de la letra para obtener resultados confiables; asimismo, aclaran si las pruebas se aplican sólo de manera individual (como en el caso de algunas de inteligencia), si pueden aplicarse de manera colectiva y con qué restricciones, o bien, de ambas formas.

Finalmente, por los **atributos que evalúan,** pueden ser clasificadas, arbitrariamente, en pruebas del área educativo-intelectual (inteligencia, rendimiento, aptitudes, percepción, desarrollo psicológico, aptitudes para el aprendizaje, memoria, lectoescritura, hábitos de estudio, lenguaje); pruebas del área clínica (personalidad, depresión, motivación, ansiedad, habilidades sociales, adaptación, trastorno por déficit de atención, autoconcepto, socialización, autocontrol, afrontamiento, asertividad) y pruebas del área de selección y orientación vocacional (aptitudes vocacionales, preferencias e intereses profesionales y vocacionales, valores). En el último capítulo de esta obra se revisarán algunas de las pruebas objetivas o psicométricas más utilizadas actualmente en México, así como una guía para su selección dentro de una batería de evaluación psicológica.

Normas de puntuación

Una vez que hemos elegido una prueba o seleccionado una batería de pruebas, de acuerdo con el objetivo que persigamos con su aplicación, debemos de tener en cuenta el tipo de puntuaciones que proporciona dicha prueba, y que son las que van a regir la interpretación. Las pruebas psicológicas referidas a la norma nos pueden proporcionar las siguientes puntuaciones derivadas: típicas o z, decatipos o estenes, eneatipos o estaninas, percentiles, puntuaciones T y, en algunos casos, edades escalares, que son las puntuaciones en las que se transforman las obtenidas directamente de la aplicación de la prueba, a las que también se les llama puntuaciones naturales, crudas o brutas; en las pruebas referidas al criterio, se trabaja con las puntuaciones obtenidas directamente o naturales y, a lo más, se les convierte en porcentajes.

En las pruebas referidas a la norma existen tres conceptos clave relacionados con las puntuaciones derivadas: normalización, estandarización y baremación. La **normalización**

se presenta cuando, al aplicar la prueba a la población meta o, más bien, a la muestra seleccionada de la población meta para obtener las puntuaciones normativas, la distribución de estas puntuaciones sigue una curva normal; aunque en sentido estricto no sea perfectamente una curva normal, si el error es mínimo, se ajustan los datos de la distribución obtenida a una curva normal.

La **estandarización** es el proceso mediante el cual se aplica una prueba a la población meta (en este caso es una muestra representativa de un país). Para que la muestra sea representativa, debe ser lo suficientemente grande (de acuerdo con el nivel de confianza determinado), y elegir a los participantes en forma totalmente aleatoria y por estratos de la población (se considera así que una prueba está estandarizada en México, en Estados Unidos, en Puerto Rico, etc.).

Llevar a cabo una estandarización es muy costoso. Por ejemplo, si quisiéramos estandarizar una prueba para niños escolarizados de seis a diez años de edad en México, tendríamos que saber primero cuántos niños de esas edades hay en el país (supongamos que 30 millones); después, elegir con qué nivel de confianza queremos que nuestra muestra represente a toda la población (p. ej., 65%, que es un nivel bajo si tomamos en cuenta que la población es muy grande, pero un nivel de confianza más alto aumentaría considerablemente el tamaño de la muestra, lo que a su vez aumentaría los costos de la estandarización), lo cual nos diría cuántos niños deberán estar incluidos en la muestra (supongamos que 5 000); a continuación, determinaríamos entonces que debemos tener 1 000 niños por cada grupo de edad (para que sea el mismo número de niños por edad evaluada), o bien, el número adecuado de los diferentes estratos de la población de esas edades (seis, siete, ocho, nueve y diez años); enseguida, la cuestión está en saber cuántos niños de esas edades se elegirán de cada uno de los 32 Estados que conforman México, para que sea una muestra estratificada; una vez hecho lo anterior, tendríamos que contar con un listado de los niños escolarizados en cada Estado para hacer una selección aleatoria. Supongamos que en el D. F. deberíamos tener 300 niños por grupo de edad, entonces tendríamos que elegir a esos 300 niños de manera totalmente aleatoria de todas las delegaciones que conforman el DF y, una vez hecho esto, localizar a cada niño y aplicarle la prueba que queremos estandarizar, lo cual también representa un gasto mayor, porque esto mismo tendríamos que hacerlo en cada uno de los 31 Estados restantes, además de pagar gastos al personal que va aplicar las pruebas, entrenarlos, etc. ¿Se da cuenta el lector por qué es tan caro estandarizar una prueba?

Debido a que estandarizar una prueba es muy costoso y, por lo tanto, la mayoría de las veces no es posible realizarla, se habla entonces de un proceso de **baremación,** en donde las puntuaciones normativas se obtienen de una muestra lo más representativa posible —según los recursos disponibles— de una determinada población, que puede ser tan pequeña como un salón de clases, o tan grande como un país; se habla entonces de que una prueba está baremada en España, en la Zona Metropolitana de la Ciudad de México, en la UNAM, en la FES Iztacala, en la Escuela Primaria “Benito Juárez”, en el 4o. grado “A” de la Escuela “Benito Juárez”, etc. Los manuales de las pruebas baremadas especifican cuántos sujetos conformaron la muestra y cómo fueron seleccionados.

Exponemos a continuación un ejemplo de baremación. En la Unidad de Evaluación Psicológica Iztacala (UEPI), perteneciente a la Facultad de Estudios Superiores Iztacala, de la Universidad Nacional Autónoma de México (UNAM), se llevó a cabo la obtención de baremos para la zona metropolitana de la Ciudad de México, de la Batería de Aptitudes para el Aprendizaje Escolar (BAPAE) de De la Cruz (2006). Esta prueba proporciona información útil sobre las aptitudes básicas de los niños que ingresan a la educación primaria, amén de que es de fácil y rápida aplicación y se aplica a infantes de seis a siete años de edad que cursan el primer grado del ciclo básico, y está formada por cinco subpruebas: Comprensión verbal (vocabulario), Aptitud numérica (conceptos cuantitativos y manejo de números), Relaciones espaciales, Constancia de forma y Orientación espacial (estas tres últimas subpruebas conforman la Aptitud perceptiva y espacial). De la Cruz parte del supuesto de que las subpruebas de la BAPAE han sido concebidas como un instrumento que ayuda a detectar a los niños que no alcanzan, en los aspectos evaluados, un nivel básico, y considera además que las aptitudes que aprecian estas subpruebas tienen un carácter de adquiridas, lo cual significa que son susceptibles de desarrollo. El objetivo de la prueba es detectar a los niños que tienen retrasos o problemas en determinados aspectos, para ayudarles a alcanzar el nivel normal mediante estrategias educativas programadas específicamente para desarrollar tales aspectos.

Sin embargo, a pesar de lo valioso de la BAPAE, en la UEPI nos cuestionábamos si los resultados obtenidos al aplicar la batería en México podrían considerarse confiables, ya que la prueba fue baremada en España; por este motivo, decidimos obtener las normas de puntuación para la zona metropolitana de la Ciudad de México. La muestra del estudio la conformaron 1 209 niños: 643 niños (53.2%) y 566 niñas (46.8%), de seis a siete años de edad, que estaban cursando el primer año de educación primaria. La muestra fue seleccionada de 45 escuelas elegidas aleatoriamente de 11 de las 16 delegaciones políticas del Distrito Federal y de cuatro de los cinco municipios del Estado de México que conforman la zona metropolitana de la Ciudad de México (estas delegaciones-municipios también fueron escogidos aleatoriamente). Con base en los registros de la Secretaría de Educación Pública, tanto del Distrito Federal como del Estado de México, de los listados de todas las escuelas de cada delegación o municipio, se seleccionaron aleatoriamente tres escuelas: dos del turno matutino y una del vespertino. En la muestra total, 1 010 niños fueron del turno matutino y 199 del vespertino.

Una vez que se ha estandarizado o baremado una prueba, las puntuaciones naturales obtenidas de la muestra de personas que va a constituir el grupo normativo se transforman en puntuaciones derivadas. Si la distribución de las puntuaciones muestra claros indicios de seguir una distribución normal, se utilizarán las típicas o *z*; si siguen una curva normal pero planocúrtica, se utilizarán los decatipos o eneatipos; si no es cualquiera de los dos casos anteriores, se utilizarán las puntuaciones percentiles; algunos autores prefieren utilizar otro tipo de puntuaciones, conocidas como puntuaciones T.

La puntuación estándar, típica o z, indica la diferencia entre una puntuación particular y

la media aritmética en términos de desviaciones estándares (DE), esto es, expresan en unidades de DE, la distancia del sujeto con respecto al promedio del grupo:

$$Z = \frac{x - Media}{DE}$$

Una de las ventajas de este tipo de puntuaciones es que se encuentran en una escala de intervalo, a diferencia de las puntuaciones percentiles que se encuentran en una escala ordinal, como veremos más adelante.

Si al aplicar una prueba a la población elegida, la distribución de las puntuaciones obtenida por los sujetos evaluados sigue una distribución normal, se aprovechan las propiedades de la curva normal para determinar las puntuaciones. Como el lector recordará, en una distribución normal la media vale 0 y la desviación estándar 1; sin embargo, para no manejar puntuaciones negativas, mediante una transformación lineal, se iguala la media con una puntuación de 100 y la desviación estándar queda de 15 o 16.

$$Z = \frac{X - 100}{15}$$

Este tipo de puntuaciones se encuentran en las pruebas que miden inteligencia, desarrollo psicológico, habilidades psicolingüísticas y percepción visoespacial.

Entre la media y una desviación estándar a la derecha y a la izquierda de la media, se encuentra 68.26% de los casos (34.13% de cada lado), entre la media y dos desviaciones estándares se encuentra 95.46% de la población evaluada y entre la media y tres desviaciones estándares se encuentra 99.73% de la población, por lo que 0.085% de la población se encuentra por arriba de tres desviaciones estándares alrededor de la media y 0.085% por debajo de tres desviaciones estándares (ver Figura 4-4). De esta manera, cuando se aplica una prueba psicológica en la población elegida y sigue una distribución normal, las puntuaciones normativas para comparar posteriormente a cada individuo evaluado con su grupo normativo son las comprendidas en el rango que queda determinado por la puntuación media obtenida y la desviación estándar, esto es, a la puntuación media se le suma y se le resta el valor de la desviación estándar y este será el rango considerado como puntuación normativa ([media + DE; media - DE]), lo cual significa que se le compara con la ejecución que realiza 68.26% de la población para determinar lo adecuado o no de su comportamiento. Claro está que, por ejemplo, si el individuo se encuentra por arriba de ese 68.26%, no sería un problema, por el contrario, una ventaja, ya que se encontraría por arriba de la norma y tendría una ejecución mejor en inteligencia, desarrollo, lenguaje o percepción que 84% de la población (el 50% de la población que se encuentra del lado izquierdo de la curva normal, más el 34.13% que se

encuentra entre la media y una DE del lado derecho de la curva); por otro lado, aquel individuo que se encontrara por debajo de esa puntuación que obtiene 68.26% de los casos, en realidad se encontraría por abajo de 84% de la población (abajo de 34.13% más 50% del lado derecho de la curva normal), lo cual es un fuerte indicador de que hay un problema. Por ejemplo, si un niño de siete años no resuelve el número de operaciones de aritmética que resuelve 84% de niños de su edad, nos hace pensar que tiene un problema de aprovechamiento y que quizá a eso se debe su bajo rendimiento académico.

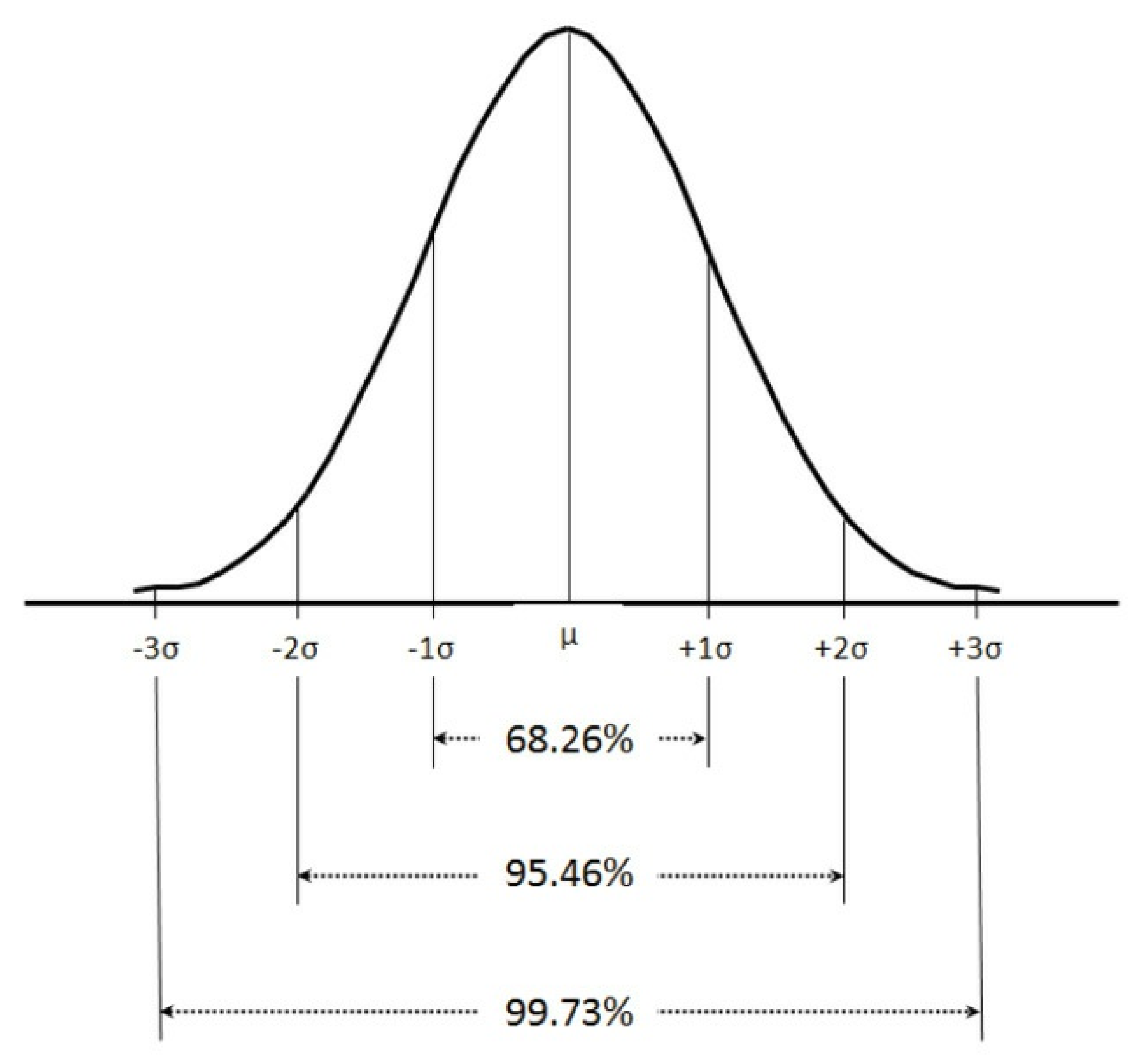


Figura 4-4. Distribución de casos en una curva normal

Las puntuaciones estándares normalizadas pueden adoptar cualquier forma conveniente, Si la puntuación estándar normalizada se multiplica por 10 y se suma o resta de 50, se convierte en una calificación T, en la que la puntuación de 50 corresponde a la media y las puntuaciones de 40 y 60 se encuentran a una DE de la media, respectivamente:

$$T = 10z – 50; T = 10z + 50$$

Es decir, puntuaciones por arriba de la norma son aquellas puntuaciones T iguales o superiores a 60; puntuaciones por debajo de la norma son puntuaciones iguales o inferiores a 40. Algunas pruebas que utilizan este tipo de puntuaciones son el MMPI-2, el Cuestionario de Ansiedad Manifiesta Revisado (CMAS-R2), la Escala Multidimensional de Asertividad (EMA) (Pruebas publicadas por El Manual Moderno).

Otras transformaciones de este tipo son los decatipos y los eneatipos. Los decatipos y eneatipos son también puntuaciones derivadas de una distribución de curva normal, sólo que en lugar de tener la forma de campana de Gauss, es una curva normal planocúrtica, es decir, aplanada, en la que en lugar de dividirse la curva en seis partes (tres a la derecha y tres a la izquierda de la media, con un tamaño igual cada fracción a una desviación estándar), se divide en 10 para los decatipos y en nueve para los eneatipos (ver Figura 4-5), por lo cual varían los porcentajes de población que se encuentran dentro de cada porción. Los decatipos o estenes son una escala de puntuaciones estándares que dividen a la distribución en 10 partes (la palabra estén es una contracción del inglés standar ten), con una media de 5.5 y una DE de 2. En los decatipos uno y 10 se encuentra 2.3% de la población, en los decatipos dos y nueve, 4.4%, en los decatipos tres y ocho, 9.2%, en los cuatro y siete, 15% y en los decatipos centrales cinco y seis, se encuentra 19.1% de la población. La interpretación toma como norma de comparación los cuatro decatipos centrales, cuatro, cinco, seis y siete, que constituirían 68.2% de la población, y considera puntuaciones extremas los decatipos uno, dos y tres del lado izquierdo de la curva (que representarían 15.9% de la población con puntuaciones más bajas) y los decatipos ocho, nueve y 10 de la derecha (15.9% con las puntuaciones más altas).

Un ejemplo de prueba que se distribuye de esta manera es la de personalidad, elaborada por R. B. Cattell, que incluye el cuestionario de 16 factores de personalidad o 16PF, el HSPQ, el CPQ y el ESPQ, que evalúan la personalidad de personas mayores de 16 años, adolescentes de 12 a 16 años, niños de 8 a 12 años y niños de 6 a 8 años, respectivamente. Al seguir las puntuaciones de esta prueba una distribución de curva normal de decatipos, Cattell interpreta los decatipos 1, 2 y 3 como el polo negativo del factor o rasgo de personalidad que se evalúa, por ejemplo: reservado, baja capacidad mental escolar, poca estabilidad emocional, sumisión, poco autocontrol, relajado; en contraste con el polo positivo, que correspondería a los decatipos 8, 9 y 10: afectuoso, alta capacidad mental escolar, emocionalmente estable y maduro, dominante, bastante autocontrol, ansiedad; y sitúa las puntuaciones intermedias: 4, 5, 6 y 7 como poseedoras del rasgo en término medio o adecuado, como lo posee 68.2% de la población y que lo demuestra en mayor o menor grado, según el contexto en el que se encuentre.

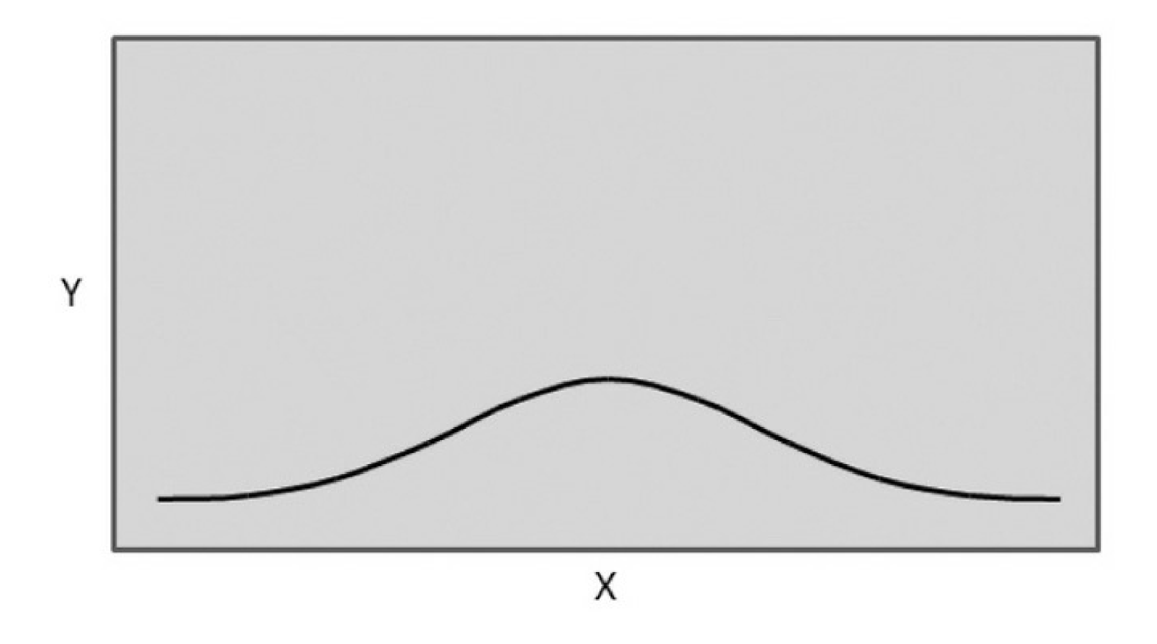


Figura 4-5. Distribución de curva normal de decatipos

Por otro lado, la escala de eneatipos o estaninas es una escala de puntuaciones de sólo un dígito, del 1 al 9, con una media de 5 y una DE aproximadamente igual a dos (la palabra estanina es una contracción del inglés *estandar nine*). En este tipo de distribución, en los eneatipos 1 y 9 se encuentra 4% de la población; en 2 y 8, 7% de la población; en 3 y 7, 12%; en 4 y 6, 17%, y en el eneatipo central se encuentra 20% de la población. Para fines normativos, se consideran como puntuaciones adecuadas a los tres eneatipos centrales 4, 5 y 6, esto es, las que obtienen 54% de la población; a los eneatipos 1, 2 y 3, como puntuaciones bajas, y a los eneatipos 7, 8 y 9, como puntuaciones altas (respectivamente, 23% de la población). Una prueba que sigue este tipo de distribución es la de Hábitos de Estudio de Pozar que, de acuerdo con el eneatipo obtenido, se interpreta como hábitos de estudio adecuados, deficientes o buenos.

Sin embargo, como no es factible que toda prueba construida se distribuya normalmente al aplicarla a la población meta –que son la mayoría–, las puntuaciones que se utilizan son los percentiles (o centiles), puntuaciones derivadas de una distribución de frecuencia acumulada y dividida en 100 partes, con el mismo número de casos cada una (a diferencia de la curva normal, donde cada parte de la curva tiene un porcentaje diferente de población). Algunos percentiles importantes son la mediana y los cuartiles primero y tercero, que corresponden, respectivamente, a los percentiles 50, 25 y 75, y son importantes porque la normatividad se establece en las puntuaciones comprendidas entre el percentil 25 y 75.

La mediana (el percentil 50) es el punto de una distribución de frecuencias acumulada que divide a la población en dos partes iguales: por arriba y por debajo de la mediana se encuentra 50% de los casos. Los cuartiles dividen la distribución de frecuencia acumulada en cuatro partes iguales, con 25% de la población cada una; así, por debajo

del cuartil uno (percentil 25) se encuentra 25% de la población que obtuvo las puntuaciones más bajas y por arriba del tercer cuartil (percentil 75), 25% de la población con las puntuaciones más altas; el 50% de la población restante se encuentra entre los cuartiles uno y tres (ver Figura 4-6).

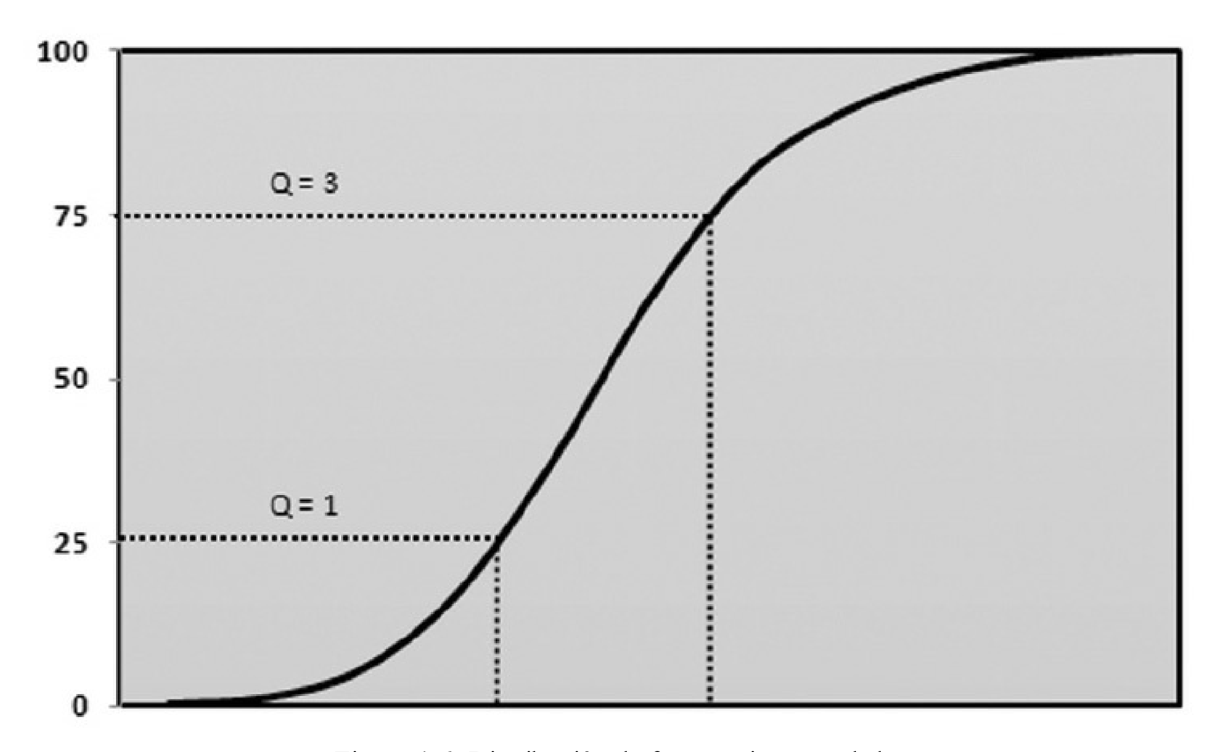


Figura 4-6. Distribución de frecuencia acumulada

A diferencia de un estadístico como la media, que es sensible a las puntuaciones extremas, una de las ventajas de utilizar las puntuaciones percentiles es que son poco sensibles a puntuaciones extremas, es decir, si alguno o algunos sujetos de una población obtienen puntuaciones muy bajas o muy altas, esto no cambia la distribución. Por ejemplo, si una muestra de cinco niños obtiene los siguientes puntajes en una prueba: 1, 2, 3, 4 y 5, en este caso la media es 3, y la mediana también es 3; sin embargo, si las puntuaciones fueran: 1, 2, 3, 4 y 100, la media ahora es 22 pero la mediana sigue siendo 3. Con esto se ilustra cómo una puntuación extrema en un conjunto de valores puede cambiar radicalmente la media, sin embargo la mediana sigue siendo la misma, porque como ya dijimos, la mediana no es sensible a puntuaciones extremas. De esta manera, con respecto a esta última media, los individuos con una puntuación de 1, 2, 3 y 4, se ven afectados, puesto que esa puntuación obtenida se aleja enormemente de la media (de 18 a 21 puntos, a diferencia de la primera media, en que la diferencia estaba en un rango de 1 a 2 puntos), pero en el caso de la mediana no sucede esto, conservan su misma posición dentro de la distribución. Esto es precisamente lo que nos proporcionan las puntuaciones percentiles, una medida de posición.

En una puntuación z, por ejemplo un CI de 80, esta puntuación nos indica la cantidad del atributo que posee el individuo evaluado; comparado con el grupo normativo decimos

que se encuentra por debajo de la norma (que es de 90 a 110). Sin embargo, una puntuación percentil se encuentra en una escala ordinal, donde la puntuación obtenida no representa la cantidad del atributo medido, sino el porcentaje de la población por arriba del cual se encuentra el sujeto evaluado. Por ejemplo, si un niño tiene un percentil de 80 en una prueba de ansiedad, ese 80 no es el porcentaje del atributo evaluado, sino que ese 80 es el porcentaje de niños de su edad a los cuales sobrepasa el sujeto evaluado, esto es, ese niño está presentando más ansiedad que el 80% de los niños de su edad; no podemos cuantificar su ansiedad, sólo sabemos que tiene más o menos del atributo evaluado comparado con los otros niños de su población.

De esta manera, si al aplicar una prueba a la población meta sus puntuaciones no se distribuyen normalmente, entonces es conveniente elaborar una distribución de frecuencias acumuladas y determinar los diferentes percentiles, tomando en cuenta que los percentiles de comparación serán el 25 y el 75, ya que entre ellos se encuentra 50% de la población con puntuaciones intermedias, es decir, la mitad de la población que se halla en medio de la distribución, y con ella se va a comparar la ejecución de los demás evaluados. Así, si un sujeto al aplicarle una prueba de inteligencia se encuentra en el percentil 65, quiere decir que se sitúa por arriba de 65% de la población en ese atributo y que a su vez 35% de la población tiene una ejecución mejor que él; si se encuentra en el percentil 85, su ejecución se encuentra por arriba de 85% de los chicos de su edad y sólo por abajo de 15% de la población, pero si su percentil es de 23, entonces esto indicaría que su ejecución se encuentra por abajo de 77% de la población normativa. Si evaluamos ansiedad, entonces encontrarse en el percentil 25 o menos no sería un problema, pero sí encontrarse en el percentil 75 o más arriba, ya que indicaría que el sujeto evaluado sobrepasa al 75% o más de la población en ansiedad. Las normas aceptadas universalmente para determinar lo adecuado o no de un atributo medido son los percentiles 25 y 75; si el evaluado obtiene una puntuación percentil entre 26 y 74, se encuentra en la norma, cerca de la norma alta de 65 a 74 y de la norma baja de 35 a 26; puntuaciones iguales o mayores a 75 serían consideradas altas e indicarían en qué porcentaje sobrepasa el evaluado a la población en el atributo evaluado (por ejemplo, inteligencia, que sería muy bueno, pero ansiedad o depresión, que serían problemáticos), y puntuaciones iguales o menores a 25 se considerarían bajas e indicarían al restarse de 100 el porcentaje por abajo del cual se encuentra el evaluado en el atributo, con respecto a la población de comparación (p. ej., inteligencia, que sería muy baja, pero ansiedad o depresión, que serían muy adecuados).

Desarrollemos un sencillo ejercicio para ejemplificar este tipo de puntuaciones, que son las más utilizadas en la actualidad por las pruebas psicológicas.

Vamos a suponer que aplicamos una prueba de Aritmética en tres diferentes escuelas (A, B y C), en la Tabla 4-1 se presenta la frecuencia absoluta o el número de alumnos que obtiene determinada calificación:

Tabla 4-1. Frecuencia de calificaciones de alumnos de tres escuelas

CALIFICACIONES	ESCUELA A	ESCUELA B	ESCUELA C

	Frecuencia	Frecuencia	Frecuencia
4	6	1	0
5	10	5	1
6	15	10	5
7	15	15	8
8	8	15	21
9	5	10	20
10	1	4	5

Para poder hacer comparaciones con puntuaciones percentiles, necesitamos calcular la frecuencia acumulada, cálculos que se presentan en la Tabla 4-2:

Tabla 4-2. Frecuencias absoluta y acumulada de calificaciones de alumnos de tres escuelas

CALIFICACIONES	ESCUELA A		ESCUELA B		ESCUELA C	
	Frec.	F.Acum	Frec.	F.Acum	Frec.	F.Acum
4	6	6	1	1	0	0
5	10	16	5	6	1	1
6	15	31	10	16	5	6
7	15	46	15	31	8	14
8	8	54	15	46	21	35
9	5	59	10	56	20	55
10	1	60	4	60	5	60

Para obtener el valor de un determinado percentil o partición, se emplea la siguiente fórmula:

$$Px = Li + \frac{\frac{N}{x} - fa}{fi}(i)$$

Por ejemplo, obtengamos el percentil 25 del grupo A. Lo primero que tenemos que hacer es localizar el intervalo donde se encuentran hasta 15 sujetos, que sería la cuarta parte de la población o percentil 25 (N/x = 60/4 = 15), ese intervalo es el de la calificación de cinco; hasta ese intervalo hay 16 sujetos (frecuencia acumulada, los seis que sacaron una calificación de cuatro y los 10 que sacaron 5). El límite real inferior de ese intervalo es 4.5 (el intervalo iría de 4.5 a 5.4); FA es la frecuencia acumulada en el intervalo anterior al intervalo en el que estamos localizando la partición (N/x), que en este caso es de seis; FI es la frecuencia absoluta del mismo intervalo, que aquí es de 10; i es el tamaño del intervalo, que en este caso es de una unidad. Con estos datos ya podemos desarrollar la fórmula:

$$P25 = 4.5 + \frac{15 - 6}{10}(1) = 4.5 + \frac{9}{10} = 4.5 + 0.9 = 5.4$$

Obtengamos ahora para el mismo grupo el percentil 75:

$$P75 = 6.5 + \frac{45-31}{15}(1) = 6.5 + \frac{14}{15} = 6.5 + 0\ 93 = 7.43$$

Si obtuviéramos los percentiles 25, 75, 50 y la media de las tres escuelas, tendríamos las siguientes puntuaciones (ver Tabla 4-3):

Tabla 4-1. Frecuencia de calificaciones de alumnos de tres escuelas

	ESCUELA A	ESCUELA B	ESCUELA C
Percentil 25	5.4	6.4	7.55
Percentil 75	7.43	8.43	9.0
Percentil 50	6.43	7.43	8.2
Media	6.47	7.4	8.05

A continuación presentamos la gráfica correspondiente de las tres escuelas (Figura 4-7):

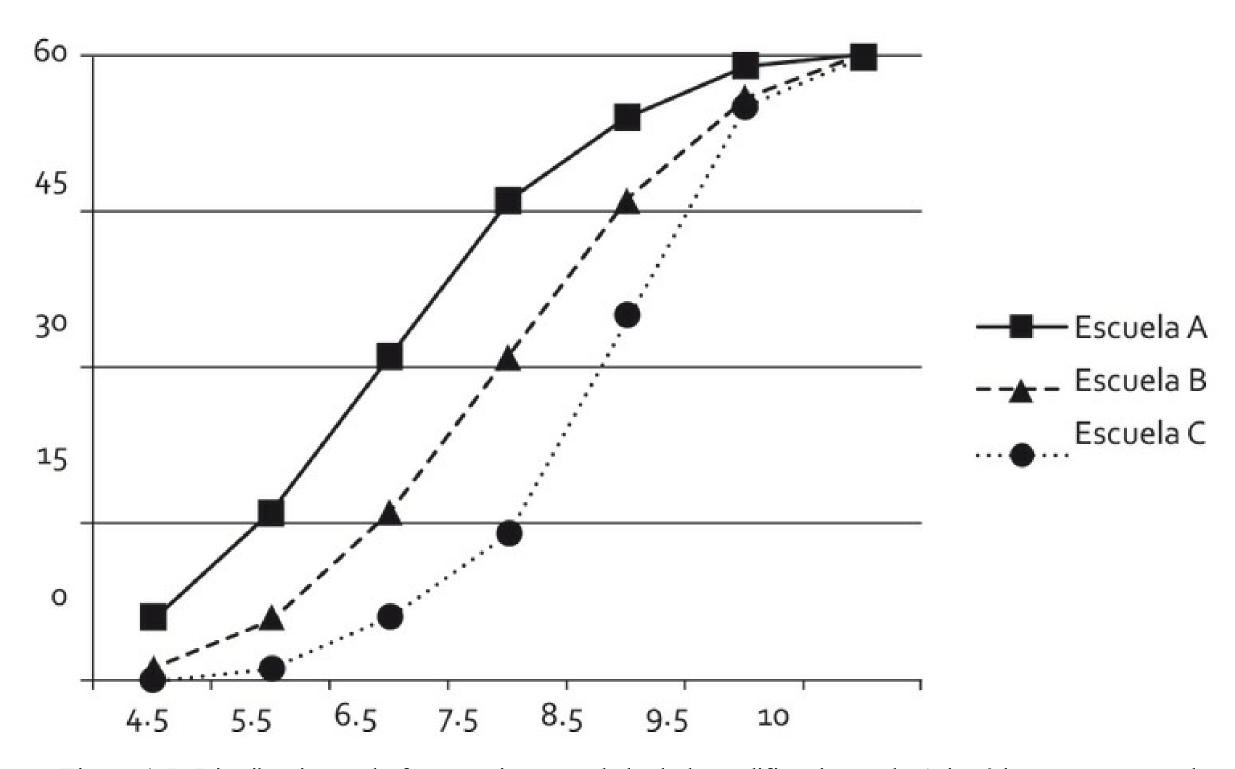


Figura 4-7. Distribuciones de frecuencia acumulada de las calificaciones de Aritmética en tres escuelas

Tanto en la tabla como en la gráfica, podemos ver relaciones interesantes. Considerando la mediana, vemos que en la Escuela A, la mitad de los alumnos (esto es, 30 alumnos), obtienen puntuaciones de 4 a 6.42 y la otra mitad, puntuaciones de 6.44 a 10; en la Escuela B, la mitad obtiene puntuaciones de cuatro a 7.42 y la otra mitad de 7.44 a 10; en la Escuela C, la mitad de alumnos obtiene puntuaciones de cuatro a 8.1 y la otra mitad puntuaciones de 8.3 a diez.

El 50% de los alumnos considerados como grupo normativos obtiene calificaciones en

la Escuela A, de 5.5 a 7.4, es decir, en este caso, las puntuaciones bajas se consideran de 5.4 o menos y las puntuaciones altas de 7.5 a diez. Para la Escuela B, las puntuaciones normativas van de 6.5 a 8.42; aquí un alumno tiene un bajo rendimiento si obtiene calificaciones inferiores a 6.5 y es un alumno de alto rendimiento con calificaciones superiores a 8.42. Para la Escuela C, un alumno se considera con puntuaciones bajas si obtiene menos de 7.56 de calificación, y para ser considerado un alumno sobresaliente debe tener una calificación mayor a nueve.

Eso es lo que hacen las puntuaciones percentiles y las puntuaciones normativas en general: comparan la actuación de cada sujeto con su grupo de referencia. Si se comparara a un individuo de un grupo con las normas de otro, se podrían falsear los datos, por ejemplo, un alumno con una puntuación de 7.5, considerada sobresaliente en la Escuela A, en la Escuela C sería una puntuación deficiente; una puntuación de 7.43 considerada justo a la mitad del grupo normativo en la Escuela B (por arriba y por abajo del 50% de la población), sería sobresaliente en la Escuela A.

Como observamos, los percentiles nos dicen la posición que ocupa un sujeto con respecto a su grupo de referencia. Un alumno por ejemplo con una puntuación de 8.5 en el grupo B, supera en el atributo medido, Aritmética, al 75% de la población. Si en lugar de Aritmética éstas fueran puntuaciones de una prueba de ansiedad, no podríamos decir que una puntuación de 8.5 en la prueba es la cantidad de ansiedad que el evaluado tiene, sino que con esa puntuación, supera al 75% de su grupo, es decir, que tiene más ansiedad que el 75% de sus pares.

Por último, las edades escalares, equivalentes, mentales o de desarrollo que establecen algunas pruebas, sobre todo las infantiles y algunas para adolescentes, es la calificación media obtenida en una prueba, por las personas de una edad cronológica determinada, es decir, proporcionan la edad promedio de ejecución del evaluado en el atributo medido. Aquí se debe tener cuidado con la interpretación y no declarar que el evaluado tiene una edad mental de X años, sino especificar que sólo es su edad de ejecución en el atributo medido. Por ejemplo, si al evaluar a un niño de siete años con una prueba de inteligencia, su CI fue de 84, su percentil de 20 y su edad escalar de 5:9, indicaría que se encuentra por debajo de la norma en cuanto a inteligencia, que su ejecución lo coloca por abajo de 80% de los niños de siete años y que su ejecución en la prueba corresponde, en promedio, a la ejecución que tendría un niño de cinco años nueve meses, esto es, tendría un atraso aproximado en inteligencia de un año tres meses; pero de ninguna manera indicaría que ese niño de siete años se comporta como un niño de cinco años nueve meses en todas las demás esferas de su vida, sino que sólo su ejecución en la prueba es la que se equipara con esa edad.

De esta manera, si tenemos un cliente o un grupo de sujetos a quienes vamos a evaluar, en primer lugar, después de la entrevista y de conocer el motivo de consulta, determinamos claramente un objetivo de evaluación, seleccionamos la o las pruebas apropiadas para lograr ese objetivo y las aplicamos y calificamos de acuerdo con el manual de la prueba. De las puntuaciones directas o crudas obtenidas al calificar la prueba, encontramos, en las tablas de los manuales de las pruebas utilizadas, las

puntuaciones derivadas que pueden ser típicas, T, decatipos, eneatipos, percentiles y/o edades escalares, con lo cual podemos hacer la interpretación de lo que significan esas puntuaciones, lo que representan para el sujeto en los atributos evaluados y a la luz de la teoría en la que está inmersa la prueba y el tipo de validez que tiene. El último paso es la entrega de un reporte de evaluación, en el que plasmemos toda la información que hemos recabado, lo cual se explicará en el Capítulo 6. En el siguiente capítulo expondremos una cuestión de suma importancia que debe tomarse en cuenta desde el inicio del proceso: la ética que rige la evaluación psicológica en general y la aplicación de pruebas en particular.

Referencias

Anastasi, A. (1988). *Psychological Testing*. New York: MacMillan Publishing Company.

De la Cruz, M. V. (2006). *Batería de Aptitudes para el Aprendizaje Escolar*. Madrid: TEA Ediciones.

Gregory, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México: El Manual Moderno.

Hoste, R. (1981). How Valid are School Examination? An Explotation into Content Validity. *British Journal of Educational Psychology*, 51, 10-22.

Kerlinger, N. F. (1975). *Investigación del comportamiento. Técnicas y metodología*. México: Interamericana.

Kirsch, I. S. y Guthrie, J. T. (1980). Construct validity of Functional Reading Tests. *Journal of Educational Measurement*, 17 (2), 81-93.

Linehan, M. M. (1980). Content Validity. Its Relevance to Behavioral Assessment. *Behavioral Assessment*, 2, 147-159.

Martínez A. R. (1996). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.

Muñiz, J. (2003). *Teoría clásica de los tests*. Madrid: Pirámide.

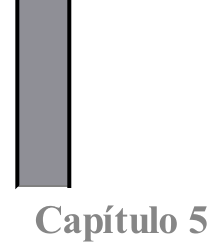
Nunnally, J. C. (1973). *Introducción a la medición psicológica.* Buenos Aires: Paidós.

Nunnally, J. C. y Bernstein, I. J. (1995). *Teoría psicométrica.* México: McGraw-Hill.

Notas

[1] Se le conoce en la literatura psicológica como TCT por sus siglas en inglés

[2] $C_m^n = \frac{n}{m!(n-m)!} = \frac{8!}{4!(4!)} = \frac{680}{24} = 70$



Capítulo 5

CONSIDERACIONES ÉTICAS EN LA EVALUACIÓN PSICOLÓGICA

La palabra ética proviene del griego *ethos*, que significa costumbres, conducta habitual, usos; el Diccionario de la Real Academia de la Lengua Española (2001) la define como la parte de la filosofía que trata de la moral y de las obligaciones del hombre; en cuanto a la ética profesional, nos dice, es el conjunto de normas morales que rigen la conducta humana. La rama de la ética que se ocupa de regular las actividades de una profesión es la deontología y se define como el conjunto ordenado de deberes y obligaciones morales que tienen los profesionales de una determinada materia. De esta manera, la ética profesional estudia las normas recogidas por la deontología profesional; la ética sugiere aquello que es deseable y condena lo que no debe hacerse, mientras que la deontología cuenta con las herramientas administrativas para garantizar que la profesión se ejerza de manera ética.

El primer código ético para psicólogos aparece en 1953 en Estados Unidos, publicado por la *Asociación Psicológica Americana* (APA), titulado *Ethical principles of psychologist* (Principios éticos para psicólogos), el cual ha tenido varias revisiones desde entonces (1992 y 2002, entre las más importantes). Los principios éticos generales de este código se resumen en: "…la competencia, la integridad y la responsabilidad profesional y científica del psicólogo, el respeto por los derechos y la dignidad humanos, la preocupación por el bienestar de los otros, y una activa responsabilidad social en la comunidad" (Frank de Verthely, 1996, pág. 14; cit. en Garaigordobil, 1998).

El conocimiento de las normas éticas permitirá comprender la responsabilidad asociada tanto con el ejercicio de la profesión en general, como en particular la actuación del psicólogo en el ámbito de la evaluación psicológica, contribuyendo así al reconocimiento tanto de sus propias capacidades con la finalidad de mejorarlas, de las limitaciones de las técnicas que emplea, así como de la comprensión de la responsabilidad social que tiene, ya que las recomendaciones sugeridas una vez terminado el proceso de evaluación, pueden alterar la vida de quienes lo consultan. En la actualidad, podemos decir que la evaluación psicológica se ha extendido bastante: las pruebas psicológicas se usan cada vez más para evaluar a personas, grupos o instituciones, tanto para tener una información

más amplia acerca de sus fortalezas y debilidades como para planear un tratamiento o intervención. De esta forma, las escuelas, las clínicas y las personas mismas, solicitan una evaluación psicológica para tomar mejores decisiones acerca de los sujetos que presentan problemas de aprendizaje, de conducta o emocionales. Sin embargo, algunas de las preocupaciones principales relacionadas con el uso de pruebas psicológicas son la invasión o intromisión en la vida privada de las personas, la violación al carácter confidencial de los resultados obtenidos y el abuso y mal uso de las pruebas; esto último tiene que ver con la aplicación de pruebas no adecuadas para el problema o la población de que se trate, y con la aplicación e interpretación de pruebas por personas incompetentes.

El objetivo de este capítulo es acotar las consideraciones éticas que se deben tomar en cuenta cuando se va a realizar un proceso de evaluación psicológica en general y cuando se utilizan pruebas como una herramienta para la toma de decisiones, con la finalidad de garantizar los derechos de las personas a las que se ofrece el servicio de evaluación psicológica. Como menciona Muñiz (2003, p. 303), en la actualidad existe una preocupación creciente por la aplicación de los avances científicos, tecnológicos y profesionales acorde con los principios éticos en todas las ramas del conocimiento, y en psicología esto es muy importante, ya que "…el mal uso de los conocimientos psicológicos, cuyo campo de aplicación son las personas directamente, encierra un potencial especial para causar daño que hay que evitar a toda costa".

Empleo de pruebas psicológicas

El uso de pruebas psicológicas, como ya revisamos en los dos primeros capítulos de este libro, a la vez que se convirtió en un hito para la evaluación de las potencialidades humanas, desafortunadamente también se pobló de mitos y controversias, sobre todo alrededor de la Segunda Guerra Mundial, etapa en la que proliferaron pruebas de todo tipo y sin las garantías científicas de su elaboración y uso, lo que originó, como se recordará, la reglamentación tanto de la construcción como del uso de las mismas. Como cualquier descubrimiento científico, las pruebas psicológicas no están libres de riesgos, de abusos y de malos usos; sin embargo, como menciona Muñiz (2003), la posibilidad de que las pruebas puedan causar daño se debe, más que a sus propiedades técnicas, al uso que psicólogos sin preparación o sin ética les den. Así, en primer lugar, hay que insistir en que las pruebas psicológicas sólo deben emplearse por psicólogos y, en segundo lugar, por psicólogos competentes y bien preparados en su selección, aplicación, calificación e interpretación.

Aunque pareciera que lo puede hacer cualquier persona, aplicar una prueba requiere de la comprensión cabal de cómo dar las instrucciones y cómo establecer el *rapport* para que los evaluados cooperen, pero sobre todo, y lo que hace esta actividad exclusiva de los psicólogos, de la interpretación de resultados a la luz de la teoría sobre la que la prueba descansa, además de establecer una relación con los demás datos que se tienen de la persona recogidos durante las sesiones de entrevista y evaluación. Sólo un psicólogo,

de acuerdo con el motivo de consulta, podría decidir qué prueba o pruebas seleccionar para obtener información relevante para la toma de decisiones, sin olvidar factores como la confiabilidad, la validez, el objetivo de la prueba y las normas de puntuación. Por ello, la venta de pruebas psicológicas está restringida y sólo se hace previa presentación de la cédula profesional; como se menciona en el *Código Deontológico del Psicólogo de España* (1988, citado en Fernández-Ballesteros, 1996) y en el *Código Ético del Psicólogo Mexicano* (Sociedad Mexicana de Psicología, 2005), el material psicológico debe estar estrictamente reservado al uso de los psicólogos, quienes se abstendrán de facilitarlos a otras personas no competentes. A este mismo respecto, Muñiz (2003) nos dice que la forma de minimizar el uso inadecuado de pruebas ha sido y sigue siendo la misma en todo el mundo: combinar la formación académica de los psicólogos con la restricción de acceso a las pruebas de personas no cualificadas. El mismo autor menciona que la sola obtención de un título de licenciatura en psicología no es garantía de poseer los conocimientos suficientes para utilizar pruebas, por lo que la formación del psicólogo se debe de complementar con cursos de especialización en evaluación psicológica.

Por otra parte, con respecto a las editoriales, Gregory (2001) menciona que es obligación de los editores de las pruebas asumir la responsabilidad de sólo publicar aquellas que hayan demostrado su alta calidad y que cuenten con las garantías científicas del instrumento, que aparecerán en el manual respectivo de las mismas, con información detallada de los datos estadísticos sobre análisis de la confiabilidad, estudios de la validez, muestras normativas, así como otros aspectos técnicos.

Es importante señalar que el psicólogo que va a aplicar una prueba debe estar familiarizado con las instrucciones, con el manejo adecuado y correcto de los materiales de la prueba, con la forma de aplicarla, con las condiciones ambientales apropiadas para ello, con el establecimiento del *rapport*, pero sobre todo, debe estar bien capacitado y haber tenido un largo entrenamiento en la práctica de calificar las pruebas y en la manera de interpretar correctamente los resultados de las personas evaluadas. Algunas recomendaciones importantes para el establecimiento de la confianza y el *rapport* dentro de un clima ético son: la primera vez que se tiene contacto con el cliente, saludarlo por su nombre, presentarse y decirle su propio nombre, así como darle un informe breve sobre el propósito de la evaluación. Una vez iniciado el proceso, el evaluador tiene que ser empático con el evaluado, sincero, establecer un clima de calidez, pero sobre todo, de respeto.

Como lo establecen los códigos del psicólogo mencionados, la autoridad profesional del psicólogo se fundamenta en su capacitación y cualificación en las tareas que desempeña, por lo que debe estar profesionalmente preparado y especializado en el uso de métodos, instrumentos y técnicas que adopte en su trabajo. A este respecto, Gregory (2001) menciona que aquellos psicólogos que tienen un entrenamiento deficiente en los instrumentos que aplican, pueden cometer errores graves de interpretación que dañen a las personas evaluadas, y aún va más allá al advertir que la utilización inepta de pruebas puede exponer al evaluador a sanciones profesionales y demandas de tipo civil.

Invasión de la intimidad, confidencialidad y consentimiento informado

El problema de la invasión de la intimidad se refiere a la posibilidad de que la aplicación de pruebas signifique una intromisión en la vida privada de las personas, es decir, que las pruebas reflejen cuestiones íntimas de la vida de los individuos que de otra manera no se sabrían. Sin embargo, habría que tomar en cuenta tres aspectos que echan por tierra esta crítica al uso de las pruebas psicológicas:

- Toda investigación que se hace sobre la conducta (ya sea científica o para beneficiar a una persona o grupo de personas) mediante pruebas u otros métodos, tales como la entrevista o la observación, supone también la posibilidad de invadir la vida privada de las personas, y en esos casos no se cuestiona el uso de tales técnicas con fines de investigación.
- Como ya hemos visto en los capítulos anteriores, la ejecución de un individuo en una prueba psicológica representa sólo una muestra de su conducta, una muestra de lo que ese individuo puede hacer en situaciones cotidianas de la vida, y no implica, por lo tanto, como menciona Anastasi (1977, p. 568), que las pruebas "…poseen poderes misteriosos para penetrar más allá de ella".
- El cliente que acude a evaluación psicológica lo hace ya sea por voluntad propia o porque es canalizado por alguna institución (escuela, clínica, etc.), pero es él finalmente quien decide acudir a evaluación y, en última instancia, tiene el derecho de decidir "…por sí mismo en qué medida compartirá con los demás sus pensamientos, sus sentimientos y los hechos de su vida personal", derecho considerado como "…esencial para asegurar la dignidad y la libertad de autodeterminación" (*Privacy and Behavioral Research*, 1967, p. 2; citado en Groth-Marnat, 1997). Como menciona la *Asociación Psicológica Americana* (APA) en sus *Ethical principles of psychologist* (1981, citado en Sattler, 1996), el psicólogo tiene la obligación de proteger la confidencialidad de la información que obtenga en el curso de la evaluación, excepto cuando el cliente o tutor haya dado su consentimiento para revelarla.

Hay que tomar en cuenta, cuando se aplican pruebas psicológicas, como ya vimos en el capítulo 3, cuál es la finalidad de la evaluación: detección, diagnóstico, descripción, selección, explicación, predicción, tratamiento o valoración e investigación. En situaciones clínicas, el cliente acude a la evaluación psicológica con la finalidad de recibir ayuda, por lo que en estos casos está dispuesto a colaborar y descubrir aspectos íntimos de su vida, de esta manera, el psicólogo no está invadiendo una vida privada; sin embargo, aún en estas circunstancias y sobre todo en aquellas no relacionadas con la clínica, como podrían ser la investigación (en cuyo caso se debe respetar el anonimato), o bien evaluaciones con propósitos institucionales como admisión o selección, se le debe informar a los evaluados los objetivos de la evaluación, en qué consistirá la misma y el empleo que se hará de los resultados obtenidos, y solicitar su permiso para llevar a cabo

el proceso, a esto se le conoce como consentimiento informado. Además, tanto el *Código Ético del Psicólogo Mexicano* como el *Código Deontológico del Psicólogo de España,* señalan que el psicólogo rechazará llevar a cabo sus servicios cuando exista la certeza de que éstos puedan ser mal usados en contra de los legítimos intereses de las personas o grupos.

Así, el primer punto a tomar en cuenta en la ética profesional del psicólogo es que una evaluación psicológica sólo se puede llevar a cabo dentro del contexto de una relación profesional claramente definida; esto significa, de acuerdo con Groth-Marnat (1997), que la naturaleza, los propósitos y las condiciones de la relación no solamente se discuten, sino que se acuerdan. De esta manera, se le debe proporcionar al cliente información sobre el tipo y la duración de la evaluación, horarios de atención, honorarios o costos económicos, la naturaleza y los límites de la confidencialidad, y toda aquella información general que pudiera ser relevante (Groth-Marnat, 1997; Sattler, 1996).

Toda actividad científica está sometida a unas normas de actuación reguladas por el contexto social en la que se desenvuelve, y la evaluación psicológica no es la excepción; como menciona Fernández-Ballesteros (1996), la evaluación psicológica no sólo tiene que ajustarse a garantías científicas, sino que, como toda actividad humana, está sujeta a ciertos principios éticos establecidos por la comunicad de referencia. Cuando llevamos a cabo un proceso de evaluación psicológica, éste debe estar justificado, es decir, la evaluación no es un fin en sí mismo, sino un medio para tener un conocimiento más amplio sobre la problemática que aqueja al paciente, con la finalidad de tomar decisiones para intentar solucionar dicha problemática; de esta manera, se le debe explicar a las personas afectadas (niños, adolescentes, padres, profesores) el uso que se hará de la información recopilada, por lo tanto, se debe obtener el consentimiento informado de las personas involucradas; como menciona Mind (1975, citado en Silva y Moro, 1994), se les debe explicar y justificar el porqué, el qué, el cómo, el dónde y el para qué de la evaluación psicológica.

Para los *Standards of Educational and Psychological Testing* (1985, citado en Gregory, 2001), el consentimiento informado implica que se ha notificado con un lenguaje comprensible a los evaluados y/o sus padres o representantes las razones para la aplicación de las pruebas, los tipos de éstas que se utilizarán, el uso indicado y la información que se dará a conocer de los resultados obtenidos y a quiénes.

Ya que toda evaluación implica una intromisión en la vida de las personas, se deben hacer explícitos los límites de la privacidad y asegurar la confidencialidad de los datos obtenidos, y, si es el caso, a quién o a quiénes se les dará a conocer dicha información y con qué fines; sólo se traspasará a terceros información pertinente en beneficio del paciente y, desde luego, como menciona Mind (op. cit.), estos terceros serán personas que puedan comprender y utilizar los resultados de la evaluación de forma correcta.

Para que podamos hablar de un consentimiento informado tanto ético como legal, se necesitan tres condiciones: divulgación de la información, competencia o capacidad de consentir y participación voluntaria (Melton, Petrila, Poythress y Slobobing, 1998; citado en Gregory, 2001). La divulgación se refiere a que el cliente reciba suficiente información

(riesgos, beneficios, difusión de informes) para que tome la decisión de aceptar la evaluación y continuar en ella; la competencia alude a la capacidad mental de la persona para dar su consentimiento, cuando esto no es posible, deberá existir un tutor legal; la participación voluntaria implica que la aceptación para realizar la evaluación se otorga de manera libre sin ninguna coacción.

El consentimiento informado incluye comunicar no sólo las razones para el uso de las pruebas, sino también la clase de datos que se obtendrán y los posibles usos que de ellos se harán; esto no significa que se le tienen que mostrar al cliente las pruebas o escalas en específico, más bien se le deben describir de manera general la naturaleza y el propósito de las pruebas; como menciona Groth-Marnat (1997), presentar el objetivo de la aplicación de pruebas de una manera simple, respetuosa y sincera, reducirá de modo significativo la posibilidad de que el cliente perciba la situación de evaluación como una invasión a su privacidad.

Es imprescindible informar a los niños y adolescentes, excepto a niños muy pequeños o con un importante deterioro intelectual, sobre el proceso diagnóstico con la finalidad de tener su consentimiento. Silva y Moro (1994, p. 719) mencionan que se debe cuidar mucho el vínculo interpersonal que se establece entre el niño o el adolescente y el evaluador, ya que puede convertirse en un vínculo personal y afectivo, facilitado porque a veces los niños y adolescentes que asisten a consulta "…se sienten inseguros, buscan apoyo o dependencia y, en algunos casos, la primera relación positiva con un adulto es la que tienen con el psicólogo…", pero si consideramos que esta relación es de corta duración, "… es fundamental preparar al niño no sólo para establecer una interrelación personal estrecha con el psicólogo, sino también para salir de ella. Es fundamental explicar al niño el proceso de evaluación evitando crear en él dependencias innecesarias".

En la entrevista inicial que se hace a los niños y adolescentes –antes de comenzar propiamente el proceso de evaluación psicológica a través de la aplicación de pruebas–, sugerimos realizar tres preguntas al finalizar la misma, estrechamente relacionadas con cuestiones éticas: ¿Sabes por qué te trajeron tus papás?, ¿Qué opinas? y ¿Estás de acuerdo con venir a esta unidad para que te evaluemos? Si el niño no sabe por qué va a ser evaluado o tiene una idea distorsionada, es el momento de informarle las razones explicándole de manera sencilla pero honesta el motivo de la evaluación. Por otra parte, es muy importante la opinión del propio niño o adolescente acerca del problema que lo aqueja; saber esto nos puede proporcionar información muy útil para saber cómo percibe su problema –si es que de hecho lo percibe–, cuáles son sus sentimientos y emociones sobre el mismo, y nos podría dar algunas pautas para la intervención. La tercera pregunta tiene que ver con el consentimiento informado, así como con la intención de obtener la colaboración del niño dentro del proceso.

Normas éticas

Los principios generales que se refieren a la práctica de la evaluación psicológica pueden resumirse, de acuerdo con la APA (Muñiz, 2003), en: competencia, integridad,

responsabilidad social y científica, respeto por los derechos y dignidad de las personas, procuración del bienestar de las personas y responsabilidad social, lo cual significa que los psicólogos deben hacer todo lo posible para tener el máximo de preparación y competencia profesional; mostrar en su relación con los usuarios de sus servicios una conducta honesta, justa y respetuosa; llevar a cabo de manera responsable sus actividades profesionales y de investigación; respetar los derechos y dignidad de todas las personas y tratar de contribuir a mejorar las condiciones de su comunidad y, en general, a mitigar el sufrimiento humano.

Hoogan (2004) considera que los códigos de ética de una determinada profesión atienden a dos propósitos fundamentales: el primero es ofrecer orientación a los profesionales sobre la conducta apropiada que deben exhibir en determinadas situaciones, así como abordar los principios que se contraponen, lo que él resume en "hacer el bien, evitar el mal y no ocasionar daño"; el segundo propósito es proteger la reputación de la profesión mediante una especie de contrato social entre la profesión y la sociedad, de esta manera, las asociaciones de profesionales contarán con procedimientos para actuar legalmente contra prácticas poco éticas.

Veamos ahora cuáles son las normas por las que se rige el *Código Ético del Psicólogo Mexicano*. Para este código, el principal fin que se persigue es la protección de los usuarios de los servicios psicológicos y es un código que regula el ejercicio profesional, científico y académico de los profesionales de la psicología; aquí sólo presentamos un breve resumen relacionado sobre todo con el ejercicio de la profesión en el área de la evaluación psicológica. El código está dividido en dos partes: **Principios generales** y **Normas de conducta**. La primera parte de *Principios generales del Código Ético del Psicólogo Mexicano* está compuesta de:

- **Bases generales:** todos los psicólogos mexicanos tienen la obligación de adherirse a este código y aplicarlo en su trabajo.
- **De los límites:** este código se aplica a toda actividad que desempeñe como parte de sus funciones académicas, científicas y profesionales, no así a las actividades personales que no tengan conexión o efecto con su rol de psicólogo.
- **De las sanciones:** éstas van desde amonestaciones hasta la suspensión de la licencia para ejercer la psicología.
- **Principios básicos** que rigen el comportamiento de los psicólogos. Estos principios son cuatro:

  Respeto a los derechos y a la dignidad de las personas:por su valor intrínseco como ser humano, toda persona tiene derecho al respeto de los demás, por tanto, debe recibir trato como tal. El psicólogo tiene la responsabilidad de respetar, proteger y fomentar el derecho de las personas a la privacidad, la autodeterminación, la libertad personal, la justicia, el consentimiento informado, la confidencialidad y el derecho de establecer y dar por terminada la relación con el psicólogo.

  Cuidado responsable: el psicólogo debe tener una preocupación activa por el

bienestar de los usuarios; sólo realizará actividades para las que recibió formación suficiente y es competente, y en las que posee conocimientos y destrezas actualizadas.

Integridad en las relaciones: el psicólogo debe mostrar en su desempeño precisión, honestidad, apertura, sinceridad, máxima objetividad y mínimo sesgo o prejuicio.

Responsabilidad hacia la sociedad y la humanidad: el psicólogo tiene responsabilidades científicas, profesionales y ciudadanas para con la sociedad y debe asegurarse de que el conocimiento que emplee será para fines benéficos.

Por otra parte, las **Normas de conducta** se dividen en tres capítulos: De la calidad del trabajo desempeñado por el psicólogo, Calidad de la valoración y/o evaluación psicológica y Calidad de las intervenciones psicológicas.

- **De la calidad del trabajo desempeñado por el psicólogo**. Este capítulo consta de 49 artículos comprendidos en cuatro apartados:

  Competencia y honestidad del psicólogo, relacionado con acreditación y actualización profesionales, no dañar a los usuarios y respetar derechos de minorías culturales, étnicas o sexuales.

  Calidad de la valoración y/o evaluación psicológica: el psicólogo que construye, desarrolla, adapta, administra o usa técnicas de evaluación psicológica, lo hace a la luz de la calidad científica, utilizando conocimientos actualizados, instrumentos que garanticen confiabilidad, validez y normas de puntuación, no promoviendo ni permitiendo su uso por personas no cualificadas.

  Calidad de las intervenciones psicológicas: llevar a cabo solamente intervenciones en las que esté capacitado.

  Calidad de la enseñanza-supervisión e investigación.

- **De las relaciones que establece el psicólogo**, con 80 artículos comprendidos en nueve apartados:

  Relación de autoridad

  Relaciones duales o múltiples

  Relaciones sexuales

  Relaciones económicas: acuerdos entre psicólogo y cliente sobre honorarios

  Relaciones con colegas y otros profesionales

  Terminación de las relaciones, que se refiere a finalizar la relación profesional cuando el usuario ya no necesita el servicio y no abandonar a los pacientes o clientes ni influir en ellos para dar por terminados sus servicios.

  Consentimiento informado: el psicólogo debe obtener el consentimiento informado utilizando un lenguaje entendible para los usuarios del servicio, si es que la persona posee la capacidad de consentir, tiene información significativa

sobre el procedimiento a seguir, expresa libremente y sin influencias indebidas su consentimiento y éste se ha documentado en forma apropiada; si no es el caso, se obtendrá el consentimiento de una persona legalmente autorizada.

Confidencialidad: el psicólogo debe respetar los derechos de confidencialidad de aquellos que le consultan, excepto cuando le obligue la ley; sólo deberá incluir información pertinente en informes escritos.

Relación con la sociedad y la humanidad.

- **De los resultados del trabajo del psicólogo**, que consta de 19 artículos y comprende dos apartados:

  Comunicación de los resultados: comunicar los resultados de manera clara y entendible a los usuarios.

  Confidencialidad de los resultados.

Estos dos últimos apartados los desarrollaremos ampliamente en el siguiente capítulo de esta obra.

Como ya mencionamos, el *Código Ético del Psicólogo Mexicano* tiene como finalidad proteger a los usuarios de los servicios psicológicos y establecer las normas que regulan el ejercicio profesional, científico y académico de los psicólogos. Independientemente del área de trabajo en que se desarrolle, es imprescindible que todo psicólogo conozca —desde el estudiante hasta el profesional— este código ético. Aquí sólo hemos presentado un breve resumen relacionado con el ejercicio de la profesión en el área de la evaluación psicológica, pero una recomendación es que lo adquieran y lo tengan siempre a la mano para que lo puedan consultar cuando sea necesario y se puedan cumplir cabalmente, como el mismo código establece, los principios que rigen el comportamiento de los psicólogos, a saber: respeto a los derechos y a la dignidad de las personas, cuidado responsable, integridad en las relaciones y responsabilidad hacia la sociedad y la humanidad.

Referencias

Anastasi, A. (1977). *Tests psicológicos*. España: Aguilar.

Diccionario de la Real Academia de la Lengua Española (2001). Recuperado de http://buscon.rae.es/draeI/

Fernández-Ballesteros, R. (1996). *Introducción a la evaluación psicológica*. Salamanca: Pirámide.

Garaigordobil, L. M. (1998). *Evaluación psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro*. España: Amarú Ediciones.

Gregory, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México: El Manual Moderno.

Groth-Marnat, G. (1997). *Handbook of Psychological Assessment*. New York: John Wilely & Sons, Inc.

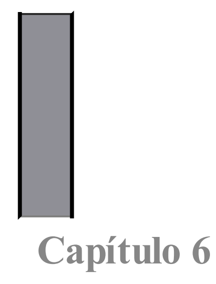
Hogan, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.

Muñiz, J. (2003). *Teoría clásica de los tests*. Madrid: Pirámide.

Sattler, J. M. (1996). *Evaluación infantil*. México: El Manual Moderno.

Silva, F. y Moro, M. (1994). Evaluación conductual en niños y adolescentes. En R. Fernández-Ballesteros (Ed.), *Evaluación conductual hoy. Un enfoque para el cambio en psicología clínica y de la salud.*

Madrid: Pirámide.
Sociedad Mexicana de Psicología (2005). *Código Ético del Psicólogo*. México: Trillas.



Capítulo 6

COMUNICACIÓN DE RESULTADOS: ELABORACIÓN DEL REPORTE DE EVALUACIÓN PSICOLÓGICA

El último paso de un proceso de evaluación psicológica, pero no menos importante, es el de la comunicación de resultados. Esta comunicación de resultados se debe entregar en un escrito en el que se plasme toda la información que hemos recabado a lo largo del proceso y que sea de utilidad tanto para el psicólogo que va a planear la intervención, si el caso lo requiere, como para los usuarios del servicio, para que puedan tomar decisiones. En principio, los resultados de la evaluación psicológica se deben comunicar a la persona evaluada; sin embargo, aquí surgen algunos dilemas éticos, ya que a veces los evaluados son niños o adolescentes que van acompañados de sus padres a la evaluación, o bien, la persona evaluada (sin importar su edad) es enviada a evaluación por la escuela, el trabajo, la clínica donde es atendida, etc. ¿Qué hacer en tales casos? El objetivo de este capítulo es dar respuesta a esta pregunta de a quién o quiénes hay que entregar los resultados, así como exponer detalladamente la forma en que se estructura por escrito un reporte de evaluación psicológica.

¿A quién se entregan los resultados de una evaluación psicológica?

Una vez que se ha terminado el proceso de evaluación psicológica, lo que procede es hacer la entrega de resultados. Si los niños tienen más de seis años, se les debe explicar los resultados de su evaluación con un lenguaje sencillo apropiado a su edad, además de, claro está, exponérselos a sus padres. Si los evaluados son adolescentes, se les explicará del mismo modo, pero se les pedirá el consentimiento para proporcionar a sus padres dichos resultados —se les hará ver las ventajas que tendría para ellos el que sus padres los conozcan— y se respetará su decisión. Cuando la persona es enviada a evaluación por alguna institución, desde el comienzo del proceso se le debe informar que, además de a ella y, si lo amerita, sus padres, el informe se dará a conocer a la institución correspondiente; en este último caso, Anastasi (1977) menciona que esta información sólo se otorgará a terceros cuando exista una necesidad genuina y legítima de poseer tal

información (como pueden ser los maestros) y, además, sólo se proporcionará la información adecuada a esa necesidad, pero desde luego, previo consentimiento del evaluado y/o de sus padres. La APA establece que la información que un psicólogo obtiene por medio de cualquier fuente sólo puede ser traspasada a un tercero con el permiso del cliente; aunque hay excepciones, como cuando los evaluados pueden dañarse a ellos mismos o a terceros. El derecho de controlar la información está claramente definido como perteneciente al cliente (Groth-Marnat, 1997).

Hay dos casos en donde es ético no comunicar los resultados al evaluado, y es cuando el psicólogo tiene evidencias de que el conocerlos empeoraría el estado emocional del paciente (p. ej., en casos de depresión, intento de suicidio, ansiedad extrema, etc.), o cuando el evaluado no puede comprenderlos, como en el caso de niños muy pequeños (menores de seis años) o con alguna discapacidad intelectual. En esos casos, se entregarían los resultados sólo a los padres y/o a la institución correspondiente.

La forma de comunicar los resultados, además de verbal, tiene que ser por escrito, y se debe hacer de forma tal que se cumpla el objetivo último de toda evaluación: ayudar a la toma decisiones; en el siguiente apartado de este capítulo se describirán las partes que debe contener un reporte de evaluación y se ilustrarán con tres ejemplos. Los informes psicológicos deben ser claros, precisos, rigurosos e inteligibles para su destinatario, expresar sus alcances y limitaciones, el grado de certidumbre que posea el evaluador, así como las técnicas utilizadas en su elaboración y los datos del profesional que lo expide. Para la APA (1981, citado en Sattler, 1996), los clientes tienen derecho a conocer los resultados de su evaluación y éstos se le deben proporcionar como base para las conclusiones y recomendaciones.

Groth-Marnat (1997) considera que la entrega de resultados de la evaluación no debe ser sólo informativa, sino que, si se proporciona de una manera adecuada, puede servir realmente para reducir la ansiedad sintomática y elevar la autoestima; asimismo, puede considerarse como una parte del proceso de intervención, ya que al psicólogo evaluador con frecuencia se le pide ayuda para tomar decisiones de vida importantes, por lo que el impacto potencial de la entrega de la información no debe ser subestimado. De esta manera, considera que los psicólogos están en una posición de poder, por lo que deben asumir la responsabilidad de la información que recibe el cliente y de las decisiones que tomará basadas en ella, y que con seguridad le acompañarán durante muchos años.

Reporte de evaluación psicológica

Como ejemplo de los elementos que debe contener un reporte de evaluación psicológica, presentamos a continuación tres de ellos: el de una niña en la pubertad que cursa la primaria con problemas de rendimiento académico; el de una adolescente de secundaria con problemas de aprovechamiento y de conducta, y el de una chica que está saliendo de la adolescencia y cursa la preparatoria con problemas emocionales a raíz de un accidente. Cabe aclarar que, para proteger la identidad de los usuarios, los nombres fueron cambiados, los antecedentes descritos son muy breves y los expedientes no son recientes.

Un reporte de evaluación debe tener los siguientes elementos: carátula, motivo de la evaluación, antecedentes, batería de pruebas aplicadas; descripción, análisis e interpretación de los resultados; actitud durante la evaluación, conclusiones, sugerencias y firma. La información de los tres primeros apartados a desarrollar: Carátula, Motivo de la evaluación y Antecedentes, se deriva de la entrevista inicial realizada tanto con el evaluado como con sus padres, por lo que antes de describirlos, acotaremos algunas cuestiones acerca de la entrevista.

La entrevista es el instrumento más empleado en todas las áreas de la psicología aplicada y también en el campo de la evaluación psicológica y su objetivo es obtener información pertinente, confiable y válida acerca del problema por el que un usuario acude a solicitar un servicio psicológico. La entrevista se caracteriza por ser una interacción verbal entre al menos dos personas, tiene objetivos prefijados, una asignación de roles bien determinados en los que el control de la situación lo tiene el entrevistador e implica además la manifestación de toda una gama de canales de comunicación humanos: verbales, no verbales (gestuales y posturales), paralingüísticos (tono, volumen, intensidad y manejo del silencio), auditivos y aún olfativos; no es un encuentro entre dos iguales, puesto que, como ya mencionábamos, hay una diferenciación de roles entre los participantes, correspondiéndole al entrevistador la responsabilidad de conducir la entrevista con la finalidad de cubrir el objetivo primordial que es el de identificar las principales áreas problemáticas del entrevistado, quien es el que ha solicitado los servicios psicológicos como consecuencia de sus propias necesidades y motivaciones.

Algunas de las ventajas de la entrevista para la evaluación psicológica son: la relación personal que se establece entre el entrevistado y el entrevistador, con el valor empático y emocional que esto conlleva y que puede facilitar posteriormente tanto el establecimiento del *rapport* como el proceso de evaluación; la flexibilidad para que el entrevistador se adapte a las peculiaridades del entrevistado de acuerdo a las necesidades del caso (retroceder sobre la marcha, avanzar, cambiar el lenguaje utilizando uno más sencillo, profundizar en algunas preguntas para obtener más información, etc.); la posibilidad de observar el comportamiento del entrevistado; la factibilidad de obtener grandes cantidades de información tanto objetiva como subjetiva y finalmente, a veces es el único medio de evaluar a las personas que difícilmente pueden ser examinadas por otros medios (Aragón, Olivares, Rodríguez y Romero, 2002).

Sattler (1996) hace algunas recomendaciones para fomentar en los entrevistadores habilidades para la entrevista:

- Escuchar de manera creativa y empática, estar libre de preocupaciones, brindar toda la atención al entrevistado; estar atento no sólo a lo que dice sino también a cómo lo dice, así como percatarse de lo que no menciona
- Observar tanto la voz como el discurso del entrevistado: sonoridad, intensidad, tono de voz, velocidad del habla, facilidad o bloqueo, espontaneidad, desviaciones del discurso, organización del mismo, vocabulario, dicción, fluidez
- Observar conductas no verbales: conducta motora, postura, expresiones faciales,

contacto visual, estados de ánimo
- Observar la apariencia personal

Ya mencionábamos que una de las ventajas de la entrevista es el establecimiento del *rapport*, el cual facilitará a su vez que todo el proceso de evaluación se lleve a cabo con éxito. El *rapport* consiste en crear una atmósfera cálida y de aceptación, su objetivo es que el entrevistado y/o evaluado se sienta cómodo, tranquilo, comprendido, a salvo, y que se comunique de manera abierta sin temor a ser juzgado o criticado. Por lo tanto, el *rapport* está basado en la confianza, el respeto y la aceptación mutuos. Sattler (1996) menciona que una forma de facilitar el establecimiento del *rapport* es que el entrevistador y/o evaluador mantenga un contacto visual apropiado, muestre una postura natural, relajada y de atención; hable de manera lenta, clara y amistosa, con un tono cálido y expresivo; que no juzgue al evaluado, que haga de la evaluación una empresa conjunta y que demuestre interés, comprensión y reconocimiento.

En resumen, el objetivo principal de una entrevista de evaluación psicológica es identificar las principales áreas problemáticas que aquejan al cliente; sus objetivos secundarios son determinar las conductas y situaciones relacionadas con las conductas problema, así como establecer el *rapport* entre el entrevistado y el entrevistador. Finalmente, no hay que olvidar que todo el proceso de evaluación, desde la entrevista hasta la entrega de resultados, tiene que ser objetivo, esto es, el psicólogo tiene que hacer a un lado sus prejuicios y puntos de vista subjetivos en el planteamiento de sus hipótesis, interpretación de resultados y recomendaciones.

La entrevista que se realiza cuando el usuario es un niño o un adolescente, se realiza en dos sesiones, en la primera se entrevista a los padres o tutores y en la segunda al propio niño o adolescente; recomendamos que cuando los padres soliciten el servicio de evaluación se les informe que en la primera sesión acudan solamente ellos, sin la presencia de su hijo, para que el psicólogo les pregunte todo lo que necesite saber acerca del problema del niño sin que él esté presente y, por otra parte, para que a su vez también ellos se sientan en confianza de expresar todo lo que consideren que corresponde a la problemática del chico; en una segunda sesión se llevará a cabo la entrevista con el evaluado, aprovechando además esta sesión como una oportunidad para establecer el *rapport* y conseguir su colaboración para el futuro proceso de evaluación. Esta es una manera de obtener la mayor información posible que nos ayude a seleccionar los instrumentos de evaluación a aplicar, no olvidando, como ya lo mencionábamos en el Capítulo 5, de asegurar a los entrevistados la confidencialidad de la información, por lo que al hacer un resumen de la entrevista en el rubro de Antecedentes, hay que describir únicamente la información que sea relevante y pertinente de acuerdo al problema y a quien va a leer el reporte de evaluación.

CARÁTULA

Es la primera hoja del reporte y debe contener los datos personales del evaluado:

nombre, dirección, teléfono, edad, fecha de nacimiento, grado de estudios, escuela a la que asiste y fecha de la evaluación.

Expediente de Aly:
Nombre: Aly …..
Edad: 12 años
Fecha de nacimiento:
Dirección:
Teléfono:
Escolaridad: se encuentra cursando el 5º año de primaria
Escuela:
Fecha de evaluación: … de 2007

Expediente de Bety:
Nombre: Bety …..
Edad: 14 años ocho meses
Fecha de nacimiento:
Escolaridad: se encuentra cursando el primer año de secundaria
Escuela:
Dirección:
Teléfono:
Fecha de evaluación: … de 2001

Expediente de Cecy:
Nombre: Cecy……
Edad: 17 años
Fecha de nacimiento:
Escolaridad: se encuentra cursando el 5º semestre de preparatoria
Escuela:
Dirección:
Teléfono:
Fecha de evaluación: …. 2002

MOTIVO DE LA EVALUACIÓN

Este apartado debe describir con mucha claridad, pero de manera breve y concisa, cuál es el motivo por el cual el usuario acude a solicitar una evaluación psicológica y, si es el caso, quién lo remite y por qué.

Expediente de Aly:
Motivo de la evaluación

La señora Dora solicita una evaluación para su sobrina Aly, ya que dice que la niña tiene problemas en la escuela porque no aprende rápido, en la boleta de calificaciones de 4º año obtuvo calificaciones bajas; además, desde hace cinco años no ve a su mamá porque se fue a Estados Unidos y la niña la extraña, y se encuentra triste.

Expediente de Bety:

Motivo de la evaluación

La señora Mónica acude a esta Unidad para solicitar que se evalúe a su hija Bety, porque presenta bajo rendimiento en la escuela y conductas de agresividad.

Expediente de Cecy:

Motivo de la evaluación

Cecy acude a esta Unidad para solicitar una evaluación psicológica debido a que hace cerca de 10 meses sufrió un accidente (la atropelló un auto) y estuvo en coma aproximadamente 12 o 13 días. Como consecuencia de ese accidente, Cecy perdió el semestre y tiene a veces fuertes dolores de cabeza, considera que se volvió más sensible y ha tenido problemas en la escuela.

ANTECEDENTES

En este punto, se describe un resumen de la historia clínica y/o entrevista con el cliente y con otros significativos (padres, maestros, etc.); la información vertida aquí debe ser pertinente y relevante a propósito del motivo de la evaluación y, de acuerdo con quien(es) va(n) a leer el reporte, confidencial.

Expediente de Aly:

Antecedentes

Aly es una niña de 12 años de edad que cursa el 5º año de primaria. Hace cinco años que no ve a su mamá porque se fue a trabajar a Estados Unidos. Anteriormente, vivía con su abuelita y su hermano Édgar, pero hace meses su hermano también pudo pasar a los Estados Unidos y ahora ella vive con su tía Dora, el esposo de ésta y sus tres primos: Gaby, de 19 años; Jasmín, de 12, y Luis, de 10.

La señora Dora reporta que Aly tiene problemas en la escuela: le cuesta trabajo aprender, en la boleta de 4º año tuvo bajas calificaciones; sin embargo, también menciona que Aly ya está avanzando porque asiste al turno vespertino, donde hay pocos niños y la maestra le pone más atención, aparte de que asiste con la psicóloga de USAER. Por otro lado, le preocupa que Aly esté triste porque extraña a su mamá y que en las noches hable sola; además, cuando la niña intentó pasar a Estados Unidos los "polleros" la trataron mal.

Por su lado, Aly menciona que sufre porque quiere estar con su mamá y su hermano; dice que se distrae con facilidad y que su abuelita le dijo que ella era de lento

aprendizaje.

Expediente de Bety:

Antecedentes

La señora Mónica, madre de Bety, solicita el servicio de evaluación para su hija, ya que refiere que es muy agresiva, no entra a clases, no pone atención, es distraída, como si estuviera preocupada; menciona que su hija no quiere ir a la escuela porque dice que sus compañeros la tratan mal. La señora señala que su hija también es agresiva en casa: como a los ocho o nueve años empezó a ser así. Bety reprobó primer año de primaria y el segundo año lo hizo dos veces; en la secundaria, el bimestre anterior tenía siete materias reprobadas y este bimestre tiene ocho materias reprobadas.

Por su parte, Bety menciona que tiene problemas en la escuela con Matemáticas, la materia más difícil según ella; se lleva bien con dos amigas, pero con los demás no porque son burlones; no pone atención en la escuela, se distrae y no se da tiempo para hacer la tarea. En su casa se lleva mejor con su mamá que con su papá, aunque ella la regaña más, y con sus hermanos, de ocho y 15 años, se pelea con frecuencia.

Expediente de Cecy:

Antecedentes

La madre de Cecy, la señora Susana, solicita el servicio de evaluación psicológica para su hija, ya que refiere que, a consecuencia de que fue atropellada por un coche hace 10 meses, Cecy, que anteriormente era una chica como todas las de su edad, se ha vuelto muy sensible, cree que todos están en contra de ella, además de que ha bajado su rendimiento escolar y sufre de fuertes dolores de cabeza. Por su parte, Cecy reporta que siente que el mundo en el que se desenvuelve no le pertenece, que le cuesta mucho trabajo estudiar, hacer amigos, llevar una buena relación con sus padres y sentirse a gusto con ella misma.

BATERÍA DE PRUEBAS APLICADAS

De acuerdo con el motivo de la evaluación, el psicólogo elegirá las pruebas que considere pertinentes. En este punto se recomienda:

a) Escribir el nombre completo de la prueba con su respectivo autor.
b) El objetivo de la prueba (esta información viene en el manual de cada prueba).
c) Una breve descripción de la misma y sobre todo, y muy importante
d) La justificación de su aplicación, esto es, por qué el psicólogo la ha elegido y qué información espera obtener de ella.

Expediente de Aly:

Batería de pruebas aplicadas

Escala de inteligencia revisada para el nivel

escolar de Wechsler WISC-RM

El objetivo de esta prueba es medir el Coeficiente Intelectual de niños entre los seis y los 16 años de edad. Esta prueba consta de dos escalas: verbal y de ejecución. La Escala Verbal a su vez consta de las subescalas información, semejanzas, aritmética, vocabulario, comprensión y retención de dígitos; la Escala de Ejecución consta de las subescalas figuras incompletas, ordenación de dibujos, diseño con cubos, composición de objetos, claves y laberintos.

Se aplicó esta prueba a Aly con la finalidad de determinar sus habilidades intelectuales y observar si son apropiadas para su edad, así como para precisar si existe relación entre éstas y sus problemas en la escuela.

Escala de ansiedad manifiesta CMAS-R de Reynolds y Richmond

El objetivo de esta prueba es valorar el nivel y la naturaleza de la ansiedad en niños y adolescentes. Esta prueba consta de cuatro subescalas: ansiedad fisiológica, inquietud-hipersensibilidad, preocupaciones sociales-concentración y mentira.

Se aplicó esta prueba a Aly para observar si, debido a los acontecimientos por los que ha pasado, presenta ansiedad y si ésta influye en su aprendizaje escolar.

Cuestionario de personalidad para niños CPQ de Porter y Cattell

El objetivo de este cuestionario es proporcionar un instrumento objetivo para el diagnóstico de la personalidad de niños en edad escolar. El CPQ evalúa 14 rasgos de personalidad de primer orden y tres de segundo orden, respectivamente: expresividad emocional, inteligencia, estabilidad emocional, excitabilidad, dominancia, entusiasmo, apego a las normas sociales, sociabilidad, emotividad, confianza en los demás, ingenuidad-astucia, seguridad en sí mismo, autocontrol y ansiedad; ajuste-ansiedad, introversión-extraversión y sensibilidad.

Se aplicó este cuestionario a Aly con la finalidad de entender un poco mejor sus características de personalidad, y sus necesidades y predisposiciones.

Expediente de Bety:

Batería de pruebas aplicadas

Escala de inteligencia revisada para el nivel escolar WISC-RM

El objetivo de esta prueba es evaluar la inteligencia o capacidad intelectual mediante cinco escalas verbales (información, semejanzas, aritmética, vocabulario y comprensión o retención de dígitos) y cinco de ejecución (figuras incompletas, ordenación de dibujos, diseño con cubos, composición de objetos y claves o laberintos).

Se aplicó esta prueba con la finalidad de observar si el problema de bajo rendimiento de Bety se debe a dificultades en su capacidad intelectual.

Inventario de hábitos de estudio de Pozar

El objetivo de esta prueba es evaluar los hábitos de estudio de los estudiantes por medio de cuatro escalas: condiciones ambientales de estudio, planificación del estudio, utilización de materiales y asimilación de contenidos.

Se aplicó esta prueba con la finalidad de determinar si Bety posee los hábitos de estudio suficientes para desempeñarse adecuadamente en la secundaria.

Cuestionario de personalidad para adolescentes
HSPQ de R. B. Cattell y M. D. Cattell

Este cuestionario evalúa 14 rasgos de personalidad de primer orden: expresividad emocional, inteligencia, estabilidad emocional, excitabilidad, dominancia, impulsividad, apego a la normas, sociabilidad, emotividad, integración social, seguridad en sí mismo, autosuficiencia, autocontrol y ansiedad, y cuatro factores de segundo orden: ansiedad, introversión-extraversión, calma-excitabilidad y dependencia-independencia.

Se aplicó este cuestionario para determinar en qué medida las características de personalidad de Bety influyen en su conducta agresiva.

Expediente de Cecy:

Batería de pruebas aplicadas

Cuestionario de 16 factores de la personalidad 16PF
de R. B. Cattell, H. W. Eber y M. Tatsuoka

El objetivo de esta prueba es evaluar la personalidad de los individuos mediante 16 factores de personalidad: expresividad emocional, inteligencia, estabilidad emocional, dominancia, impulsividad, apego a la normas, sociabilidad, emotividad, confianza, actitud cognitiva, sutileza, seguridad en sí mismo, conservadurismo, autosuficiencia, autocontrol y ansiedad, y cuatro factores de segundo orden: introversión-extraversión, ansiedad, susceptibilidad-tenacidad y dependencia-independencia.

Se aplicó esta prueba con la finalidad de conocer los rasgos dominantes de la personalidad de Cecy.

Cuestionario de adaptación para adolescentes de H. M. Bell

El objetivo de esta prueba es evaluar la adaptación de los adolescentes en cuatro áreas: familiar, salud, social y emocional.

Se aplicó esta prueba con el propósito de conocer en qué áreas se encuentra bien adaptada Cecy y en cuáles se encuentra desadaptada.

Escalas Magallanes de adaptación de Manuel García Pérez y Ángela Magaz Lago

Esta prueba tiene como finalidad evaluar la adaptación familiar (hacia el padre y la madre), la adaptación escolar y la adaptación personal de adolescentes

entre 12 y 18 años de edad.

Se aplicó esta prueba para complementar la información con la prueba de adaptación de Bell, ya que esta escala evalúa además tanto la adaptación escolar (donde Cecy reporta problemas) como la personal, la cual está muy relacionada con la adaptación emocional y la autoestima.

Inventario de hábitos de estudio de F. F. Pozar

Esta prueba tiene como objetivo evaluar los hábitos de trabajo y estudio de los escolares mediante cuatro escalas fundamentales y una adicional: condiciones ambientales de estudio, planificación del estudio, utilización de materiales, asimilación de contenidos y sinceridad.

Se aplicó esta prueba para conocer en qué áreas de estudio se encuentran deficiencias para así tomar medidas correctivas.

Inventario de ansiedad estado-rasgo de C. D. Spielberger y R. Díaz-Guerrero

El objetivo de esta prueba es evaluar si el individuo presenta ansiedad y de qué tipo, ansiedad estado o ansiedad rasgo.

Se aplicó esta prueba para determinar con mayor certeza la ansiedad que Cecy presenta, ya que en la prueba de personalidad obtuvo puntuaciones extremas.

DESCRIPCIÓN, ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Una vez aplicadas y calificadas las pruebas elegidas, se debe:

a) Describir la ejecución del niño en cada una y proporcionar los valores obtenidos.

b) Analizar esa información con respecto a las normas de la población de comparación.

c) Interpretar los resultados en relación con el motivo de la consulta, con lo observado en las sesiones de evaluación y con la información previa y las entrevistas.

Expediente de Aly:

Descripción, análisis e interpretación de resultados

Escala de Inteligencia WISC-RM

A continuación, se presenta una tabla con los resultados de Aly en la Escala Verbal:

Subescala	Puntuación normalizada	Edad escalar
Información	7	8:2
Semejanzas	7	8:10
Aritmética	7	6:10
Vocabulario	10	7:2
Comprensión	21	> 16:10
Retención de dígitos	5	6:2

De manera general, como podemos observar, en todas las subescalas —excepto comprensión— Aly obtiene puntuaciones bajas. En información, que mide los conocimientos que se han aprendido tanto de su medio como de la escuela, tiene un atraso de cuatro años; en semejanzas, la cual mide formación de conceptos, presenta un atraso de tres años; en aritmética, que evalúa razonamiento numérico y cálculo mental, tiene un atraso de cinco años; en vocabulario, la cual mide riqueza de ideas y expresión verbal, presenta un atraso de cinco años; en retención de dígitos, que evalúa concentración y memoria a corto plazo, tiene un atraso de seis años; sin embargo, de manera sorprendente, en la subescala de comprensión, que evalúa comprensión verbal, juicio moral, sentido común, empleo de conocimientos prácticos, razonamiento y uso de experiencias, presenta una puntuación alta, con una edad escalar mayor a 16 años 10 meses.

La suma de puntuaciones normalizadas nos da un CI Verbal de 81, puntuación que se encuentra por debajo de la norma para los niños de su edad.

En seguida presentamos los resultados de la Escala de Ejecución:

Subescala	Puntuación normalizada	Edad escalar
Figuras incompletas	12	8:10
Ordenación de dibujos	6	8:2
Diseño con cubos	4	< 6:2
Composición de objetos	23	< 6:2
Claves	18	< 8:2
Laberintos	0	< 8:2

En la Escala de Ejecución, en todas las subescalas, Aly se encuentra por debajo de la norma. En figuras incompletas, que mide discriminación visual, Aly tiene un atraso de tres años; en ordenación de dibujos, que evalúa organización perceptual y planeación de situaciones consecutivas, presenta un atraso de cuatro años; en diseño con cubos, la cual mide organización perceptual, capacidad de análisis y síntesis, coordinación visomotora y visualización espacial, tiene un atraso de más de seis años; en composición de objetos, que evalúa organización perceptual, anticipación visual de las relaciones entre las partes, y el todo y relaciones espaciales, también presenta un atraso de más de seis años; en claves, que mide destreza motora, velocidad psicomotriz, memoria a corto plazo, recuerdo visual, capacidad de imitación y concentración, tiene un atraso de más de cuatro años; finalmente, en laberintos, la cual evalúa organización perceptual, capacidad de planeación y previsión, coordinación visomotora, atención, concentración y velocidad, no tuvo ningún acierto, por lo que tiene un atraso de más de cuatro años.

La suma de puntuaciones normalizadas nos da un CI de Ejecución de 58, puntuación que se encuentra muy por debajo de la de los niños de su edad.

Al sumar las puntuaciones de ambas escalas, la Verbal y la de Ejecución, se obtiene un CI total de 65, puntuación que se halla muy por debajo de la de los niños de su edad.

Escala de ansiedad manifiesta CMAS-R

Aly presentó puntuaciones altas en todas las escalas de ansiedad: en ansiedad fisiológica, donde se encuentra por arriba de 91% de las chicas de su edad, reporta que le cuesta trabajo tomar decisiones, se enoja con facilidad, le cuesta trabajo dormir y se mueve mucho en su asiento; en inquietud-hipersensibilidad, se halla por arriba de 99% de las chicas de su edad; reporta que es muy nerviosa, le preocupan muchas cosas, tiene miedo, es muy sentimental, se siente mal cuando se enojan con ella y se preocupa al irse a dormir; en preocupaciones sociales-concentración, Aly también se halla por arriba de 99% de las niñas de su edad; reporta que las cosas le parecen difíciles, le parece que a los demás no les gusta cómo hace las cosas y que están en contra de ella, se siente sola, infeliz y le cuesta trabajo concentrarse. De esta manera, Aly se puede considerar como una niña con bastante ansiedad, ya que su puntuación total la coloca por arriba de 99% de la población de las niñas de su edad.

Cuestionario de personalidad CPQ

En los factores de primer orden: excitabilidad, dominancia, entusiasmo, confianza en los demás, seguridad en sí mismo y autocontrol, Aly se encuentra dentro de la norma,

como la mayoría de las chicas de su edad, al igual que en los tres factores de segundo orden: ansiedad, introversión-extraversión y sensibilidad.

Los factores en los cuales Aly sí obtiene puntuaciones significativas son: expresividad emocional, inteligencia, estabilidad emocional, apego a las normas sociales, sociabilidad, emotividad, ingenuidad-astucia y ansiedad.

De acuerdo con sus resultados, Aly es una niña con habilidades intelectuales bajas, con lento aprendizaje y comprensión, lo cual se corrobora con la ejecución que tuvo en la prueba de inteligencia.

Aly se presenta como una niña reservada y alejada, así como apacible y confiada de las demás personas. En el aspecto emocional, es poco estable y madura, con poca tolerancia a la frustración y propensa a perder el control emocional, con poco apego a las normas sociales y desatenta con las reglas que se le imponen; se presenta también como sensible y emotiva, dependiente y temerosa.

Expediente de Bety:

Descripcion, análisis e interpretación de resultados

Escala de inteligencia revisada para el nivel escolar WISC-RM

Bety obtuvo en la Escala Verbal un CI de 79, considerado por debajo de lo normal para las chicas de su edad; en la Escala de Ejecución, alcanzó un CI de 94, estimado como normal, y en la Escala Total, un CI de 85, que se considera por debajo de la norma de las chicas de su edad. A continuación se presenta una tabla donde se muestran las edades escalares alcanzadas por Bety en cada una de las subpruebas de esta escala:

Escala Verbal	Edad escalar	Escala de Ejecución	Edad escalar
Información	7 años 10 meses	Figuras incompletas	11 años 6 meses
Semejanzas	10 años 10 meses	Ordenación de dibujos	7 años 6 meses
Aritmética	6 años 2 meses	Diseño con cubos	12 años 2 meses
Vocabulario	11 años 2 meses	Composición de objetos	13 años 6 meses
Comprensión	15 años 2 meses	Claves	14 años 10 meses
Retención de dígitos	15 años 2 meses	Laberintos	15 años 2 meses

Como se observa, en la Escala Verbal, Bety se encuentra ligeramente arriba de lo esperado en comprensión y retención de dígitos, lo cual quiere decir que comprende adecuadamente el lenguaje y tiene buena memoria; sin embargo, tiene un atraso de dos años en vocabulario, lo cual corrobora que su lenguaje es muy pobre; un atraso de tres años en semejanzas, que tiene que ver con formación de conceptos, cuestión básica para el aprendizaje escolar; de seis años en información, que se refiere a los conocimientos escolarizados que debería poseer de acuerdo con el grado escolar en que se encuentra, y un atraso de ocho años en Aritmética, repertorio donde Bety está muy atrasada, lo cual nos muestra el porqué de sus dificultades en matemáticas, ya que no domina las cuestiones básicas de suma y resta.

Con respecto a la Escala de Ejecución, su desempeño global es adecuado, pero se observa un ligero atraso en composición de objetos, que tiene que ver con armado de

rompecabezas; un atraso de dos años en diseño con cubos, que se relaciona con cuestiones de abstracción; de tres años en figuras incompletas, que se relaciona con repertorios de discriminación y percepción visual, y un atraso de siete años en ordenación de dibujos, que tiene que ver con lógica; no obstante, en claves, que mide procesos de memoria, se encuentra en la norma, y en laberintos se encuentra ligeramente arriba de lo esperado.

Inventario de hábitos de estudio de Pozar

En cuanto al Inventario de hábitos de estudio, Bety refiere que sus condiciones ambientales físicas y personales son adecuadas, al igual que la planificación que ella hace del estudio y la forma en que utiliza sus apuntes y sus libros; sin embargo, refiere tener resultados no satisfactorios en cuanto a la asimilación de contenidos, esto es, al aprendizaje de lo que estudia.

Cuestionario de personalidad para adolescentes HSPQ

Finalmente, en el Inventario de personalidad, Bety se caracteriza por ser una chica reservada y no responder de manera favorable a la actuación de los profesores y la situación escolar; es seria y taciturna, se autodesaprueba y probablemente necesite que en casa sean más afectuosos y cariñosos con ella; es individualista y reprimida, no le gusta actuar en grupo, en suma, se caracteriza por ser una chica introvertida; sin embargo, demuestra tener estabilidad emocional, madurez y sabe enfrentar la realidad.

Expediente de Cecy:

Descripcion, análisis e interpretación de resultados

Cuestionario de 16 factores de la personalidad

En los factores de inteligencia, impulsividad, apego a las normas sociales, emotividad, conservadurismo y autosuficiencia, Cecy se encuentra en la norma, pues tiene estas características de personalidad como la mayoría de los adolescentes de su edad.

En cuanto a sus características más sobresalientes, es una chica muy reservada y discreta, tímida y reprimida, retraída y propensa a amargarse; presenta inestabilidad emocional, pues se perturba con facilidad, se preocupa mucho ante las frustraciones y evade responsabilidades con frecuencia; es dominante, agresiva y rebelde; desconfía de las demás personas, es suspicaz y dogmática, esto es, exige que la gente acepte la responsabilidad de sus errores; es práctica y objetiva, con preocupaciones reales; se involucra de manera fácil afectuosa y emocionalmente con las demás personas; es una chica insegura de sí misma, se autorrecrimina y autoculpa, aprehensiva, atormentada, ansiosa, se deprime y llora con facilidad, y es sensible a la aprobación o desaprobación de los demás; posee muy poco autocontrol y presenta un alto grado de ansiedad, lo cual se corroboró al analizar los factores de segundo orden, en donde también presenta ansiedad extrema.

Cuestionario de adaptación para adolescentes

Este cuestionario proporciona cinco categorías de adaptación: excelente, buena, normal, no satisfactoria y mala. Los resultados de Cecy son: en las áreas familiar, salud y social, mala, y en el área emocional, no satisfactoria.

En el área familiar, Cecy ha sentido deseos de marcharse de su casa, cree que ha decepcionado a sus padres; percibe a su madre como dominante; siente que sus padres la obligan a obedecer, que le riñen con frecuencia por su mala conducta y en ocasiones le producen miedo, le dicen que no les gustan las amigas que tiene, critican su aspecto exterior y la hacen sentir molesta. Piensa que sus padres no se han dado cuenta de que ya es mayor y que no deberían tratarla como niña, siente que ellos no la comprenden; está en desacuerdo con su forma de pensar. Sus padres riñen con frecuencia, se ponen coléricos, por lo que cree que en su casa debería haber más comprensión y afecto, refiere que se pelea frecuentemente con sus hermanos.

Con respecto al área de la salud, Cecy refiere que se resfría, tiene catarros, dificultades para respirar y sufre de anginas o laringitis con facilidad, reporta que tarda bastante tiempo para curarse de resfriados; de niña era enfermiza, tiene sus ojos sensibles a la luz y usa lentes; tiene a menudo dolores de cabeza y siente pinchazos en ésta, con frecuencia se encuentra cansada al finalizar el día y cuando se levanta por la mañana. Tuvo una lesión significativa como consecuencia de un accidente y fue sometida a una operación quirúrgica importante, por lo cual ha necesitado ver a los médicos con bastante frecuencia; se siente a menudo molesta por perturbaciones en el funcionamiento de su estómago o de sus intestinos, siente vértigos o ausencias con regularidad y náuseas, vómitos o diarreas y padece de estreñimiento; se sienta frecuentemente a la mesa sin apetito y reporta tener alguna enfermedad del corazón, los pulmones o riñones.

En el área social, a Cecy no le gusta ir a sitios donde hay muchas personas ni a fiestas ni reuniones sociales; le resulta difícil y se pone nerviosa al hablar o contestar sobre algún tema ante sus compañeros de clase o en público; le es difícil conversar con alguien que le acaban de presentar o que apenas conoce, pues se considera una persona tímida y no se siente muy segura de sí misma, no hace amigas con facilidad y, en general, le resulta difícil relacionarse con las demás personas.

Finalmente, en el área emocional, Cecy sueña despierta con frecuencia y se abstrae, pierde a menudo las cosas por no fijarse dónde las deja; muchas veces se siente sola aún cuando se encuentre con otras personas; se desanima con facilidad, y una de las cosas que lo ocasiona son las calificaciones que saca en la escuela, piensa a menudo que sus proyectos le saldrán mal y se preocupa por ello, se lamenta de las cosas que hace. Le gustaría sentirse feliz y optimista como parecen estarlo otras personas de su edad y piensa que es una persona poco feliz. Se enfada con facilidad, le disgusta su aspecto exterior y se ruboriza con facilidad.

Escalas Magallanes de adaptación

En esta prueba, Cecy mostró una adaptación familiar muy baja, tanto hacia su padre como hacia su madre. En cuanto a su padre, Cecy percibe que su papá pocas veces: la acepta como es, se interesa por sus problemas, le gustan sus ideas y las respeta, aunque no esté de acuerdo, y valora las cosas positivas que ella hace; y que a su vez, a ella casi nunca: le gusta la forma de pensar de su papá, las mismas cosas que a él, e intenta que su papá esté contenta con ella. Con respecto a su madre, Cecy considera también que rara vez le gustan las mismas cosas que a ella, su forma de pensar, y que pocas veces su

mamá respeta sus gustos e ideas, aunque no los comparta, y hace comentarios positivos de su comportamiento en casa.

En el aspecto escolar, Cecy está bien adaptada a sus profesores, pero en lo que se refiere a sus compañeros y a la escuela, su adaptación es baja. Cecy menciona que en general no le gustan los compañeros que tiene porque éstos no la defienden si alguien la critica. Reporta que no siempre realiza las tareas escolares ni presta atención a sus profesores, cree que no siempre aprende mucho asistiendo a clases.

Finalmente, a nivel personal, la adaptación de Cecy es muy baja. A menudo siente que tiene muchos problemas consigo misma, le preocupa lo que otros piensan de ella, si la critican siente vergüenza; le cuesta trabajo tomar decisiones, le dan miedo las contrariedades y dificultades, y tiene poca iniciativa y se siente mal si no consigue lo que desea. Cecy percibe que muchas veces se hace reproches por las cosas que hace, se encuentra nerviosa sin saber por qué; le gustaría morirse, está triste y aburrida, se siente mal con ella misma cuando se equivoca y piensa que es un desastre y que todo lo hace mal; siempre se ve diferente a los demás, siente ira o agresividad sin motivo aparente, le cuesta expresar lo que siente y tiene tremendas luchas con ella misma.

Inventario de hábitos de estudio

En este inventario, Cecy obtuvo una puntuación adecuada en cuanto a la utilización de materiales y una puntuación buena en sinceridad, pero en las áreas de condiciones ambientales y planificación de estudio, y asimilación de contenidos su puntuación fue no satisfactoria.

En condiciones ambientales de estudio, Cecy reporta no tener voluntad firme para estudiar y no lo hace en buenas condiciones físicas; no sabe tomar apuntes, no repasa sus exámenes antes de entregarlos, se desanima y deja lo que está estudiando si encuentra alguna dificultad que no comprende. Ni en su casa ni en la escuela la animan en materia de estudio.

En planificación del estudio, Cecy señala que deja para última hora la preparación de la mayor parte de sus materias; no ha logrado hacer del estudio un verdadero hábito, pues pierde tiempo cuando estudia por no preparar de antemano lo que necesita, no termina la tarea de trabajo y estudio que se asigna, no distribuye su tiempo de estudio entre todas sus materias ni estudia a diario.

En cuanto a asimilación de contenidos, Cecy no tiene confianza en su memoria, siente que los profesores no la tratan con justicia según la medida de sus esfuerzos, por lo tanto, repite mecánicamente cuando estudia, intenta aprender de memoria y no relaciona lo que trata de aprender con lo que ya sabe.

Inventario de ansiedad estado-rasgo

Cecy presentó en ambas escalas un alto grado de ansiedad. En ansiedad estado, la tensión, ansiedad y aprensión que presenta se consideran muy altas, ya que se encuentra ubicada en el percentil 98, esto es, su ansiedad se encuentra por arriba de 98 por ciento de la población. Por otra parte, su ansiedad rasgo —forma en que Cecy percibe el mundo que le rodea— también es alta, pues se halla en el percentil 89,es decir, tiene una forma de vida más ansiosa que el 89 por ciento de la población.

ACTITUD DURANTE LA EVALUACIÓN

Aquí se describen aquellos aspectos relevantes relacionados con el proceso de evaluación y el motivo de la consulta que tuvieron lugar desde el primer momento en que se contactó con el evaluado hasta la terminación de la evaluación.

Expediente de Aly:

Actitud durante la evaluación

Durante el proceso de evaluación, Aly se mostró cooperativa, atenta, interesada en lo que se le decía y lo que se le pedía; sin embargo, también se mostraba poco segura de sí misma y de sus respuestas, porque con frecuencia interrogaba si lo había hecho bien y pedía perdón de manera constante, y le preguntaba a la evaluadora si estaba enojada.

Expediente de Bety:

Actitud durante la evaluación

Durante el proceso de evaluación, Bety se mostró atenta, cooperativa e interesada.

Expediente de Cecy:

Actitud durante la evaluación

Durante el proceso de evaluación, Cecy se mostró cooperativa al contestar las diferentes pruebas que se le aplicaron.

CONCLUSIONES

Se trata de integrar toda la información obtenida durante el proceso —los diferentes resultados obtenidos en las pruebas aplicadas, el motivo de la evaluación y la entrevista— para ayudar en la toma de decisiones.

Expediente de Aly:

Conclusiones

De los resultados obtenidos, podemos observar que las habilidades intelectuales de Aly se encuentran muy por debajo de los niños de su edad, con un CI Verbal de 81, un CI de Ejecución de 58 y un CI Total de 65, los cuales están por debajo de la norma esperada. En promedio, Aly tiene un atraso de cuatro años y medio, es decir, su ejecución en la prueba de inteligencia es equiparable a la de una niña de siete años y medio, lo cual explica sus problemas y dificultades para aprender y enfrentar las demandas de la escuela, situación que se corrobora también en el factor de inteligencia de la prueba de personalidad, donde su ejecución la cataloga como una niña que presenta lento aprendizaje. Sin embargo, hay que hacer notar que tiene una habilidad sorprendente en comprensión verbal.

Por otro lado, con respecto a su personalidad, Aly se presenta como poco estable emocionalmente, con escasa tolerancia a la frustración y tendiente a perder el control, con poco apego a las normas sociales y desatenta con las reglas, así como sensible y emotiva, dependiente y temerosa. Aly también se caracteriza como una niña ansiosa, nerviosa, con dificultades para concentrarse y dormir, con muchas preocupaciones y miedos, siente que a los demás no les gusta cómo hace las cosas y que están en contra de ella; además de sentirse sola e infeliz. Quizá estas características de personalidad que presenta y su ansiedad se deban a los acontecimientos que ha vivido últimamente, como el intentar pasar a los Estados Unidos de manera ilegal y la separación de su madre y su hermano, lo que, desde luego, también contribuye a su bajo aprendizaje escolar.

Expediente de Bety:

Conclusiones

De los resultados obtenidos, podemos concluir que la inteligencia de Bety se encuentra por debajo de lo esperado para las chicas de su edad, y sus hábitos de estudio, aunque ella refiere que son buenos, podrían mejorarse. El bajo puntaje de Bety en la Escala de Inteligencia se debe más que nada a las deficiencias que ha arrastrado desde la primaria, las cuales no le han permitido asimilar los conocimientos actuales. Finalmente, con respecto a su personalidad, es introvertida, reservada e individualista.

Expediente de Cecy:

Conclusiones

De los resultados de la evaluación, podemos deducir que Cecy es una chica tímida y retraída, y presenta baja autoestima, lo cual se denota en la evaluación de la personalidad, en su baja adaptación social y en la insatisfactoria adaptación emocional; por otra parte, su adaptación familiar, tanto a su padre como a su madre, es baja, al igual que su estado de salud, donde presenta muchos problemas. Su adaptación escolar, en lo que se refiere a sus compañeros y a la escuela, tampoco es adecuada, lo que se refleja también en sus hábitos de estudio, donde obtuvo puntajes no satisfactorios en condiciones ambientales de estudio, planificación del estudio y asimilación de contenidos. Finalmente, tanto en el cuestionario de personalidad como en el inventario de ansiedad, Cecy presentó un alto grado de ansiedad.

SUGERENCIAS

Este último punto tiene que ver con la toma de decisiones. Basadas en la información proporcionada en el reporte, se establecen sugerencias de intervención o tratamiento a corto y largo plazo que puedan ayudar al sujeto en la resolución del problema encontrado.

Expediente de Aly:

Sugerencias

Se sugiere que Aly continúe con el apoyo que le brinda la psicóloga de USAER, a quien deberán mostrársele estos resultados para que pueda planear de mejor manera la regularización psicopedagógica; también se recomienda que, de ser posible, en casa le ayuden sus dos primas y su tía en la realización de las tareas escolares y le expliquen lo que no entiende.

Por otro lado, dada la situación en la que se encuentra Aly, se sugiere también que se le brinde mucho apoyo emocional y afecto. Asimismo, se recomienda que acuda a terapia psicológica para controlar su ansiedad, sus miedos y elevar la seguridad en sí misma y su autoestima.

- **Expediente de Bety:**

Sugerencias

Se sugiere que Bety intente mejorar sus hábitos de estudio para que concluya la secundaria y estudie alguna carrera corta que le sirva para que sea autosuficiente en el futuro. Se recomienda también que en su casa le presten más atención y traten de ser afectuosos y cariñosos con ella para que adquiera mayor seguridad. Se pide a los padres que acudan a orientación psicológica para que les enseñen cómo relacionarse de manera armoniosa con Bety y les ayuden a establecer límites para controlar la agresividad de la chica.

- **Expediente de Cecy:**

Sugerencias

Se recomienda que Cecy:

Asista a una terapia psicológica que le ayude a adaptarse mejor a su familia, a sus compañeros y a ella misma, y también que le enseñen técnicas para manejar la ansiedad que presenta.

Procure buscar y asistir a algún curso o taller sobre autoestima.

Intente mejorar sus hábitos de estudio.

Que hable con sus padres y hermanos y les diga sus sentimientos con el propósito de que exista mejor comunicación y empiecen a mejorar las relaciones familiares; que procuren que haya más armonía en la familia y le brinden a Cecy más apoyo, comprensión y tolerancia.

FIRMA

Al final del reporte, con el objetivo de darle al proceso la legalidad requerida, ya que puede ser probable que lo lea otro colega psicólogo, o bien, otro profesionista (médico, maestro, trabajadora social, juez, etc.), se debe firmar y poner todos los datos: nombre completo del evaluador, firma, número de Cédula Profesional y adscripción de la institución. Ejemplo:

Evaluadora
Mtra. Laura Edna Aragón Borja
Jefa de la Unidad de Evaluación Psicológica
Facultad de Estudios Superiores Iztacala UNAM
Cédula Profesional 651770

Hagamos un resumen del proceso de evaluación psicológica: después de haber realizado una entrevista a una persona y conocer el motivo de consulta, determinamos un objetivo de evaluación, seleccionamos la o las pruebas apropiadas para lograr ese objetivo —desde luego, pruebas confiables y válidas—, obtenemos el consentimiento informado de la persona y/o sus tutores, aplicamos y calificamos las pruebas de acuerdo con el manual de la prueba, sin cometer errores. Después de las puntuaciones directas o crudas que resultan de calificar la prueba, obtenemos de los manuales de las pruebas las puntuaciones derivadas (típicas, T, decatipos, eneatipos, percentiles y-o edades escalares), con lo cual podemos hacer la interpretación de lo que significan esas puntuaciones, lo que representan para el sujeto en los atributos evaluados y a la luz de la teoría en la que está inmersa la prueba y el tipo de validez que tiene.

La última parte del proceso de evaluación es la correcta utilización de los resultados, aspecto central en la práctica. La comunicación de los resultados se lleva a cabo por medio de la entrega de un reporte de evaluación, en el que plasmemos toda la información que hemos recabado y que sea de utilidad para nosotros, para el psicólogo que va a planear la intervención y para el usuario para que pueda tomar decisiones.

Finalmente, antes de cerrar este capítulo de la obra, queremos hacer dos recomendaciones: La primera es que, además de utilizar un lenguaje claro y sencillo para que el usuario entienda los resultados, en la medida de lo posible y sin dejar de ser honestos y precisos, hay que evitar etiquetar o dar categorías diagnósticas que puedan resultar incómodas, desmoralizantes, peyorativas o que puedan producir angustia o aflicción, pues pueden tener algún impacto negativo en la persona evaluada. Por ejemplo, en lugar de señalar que el niño es “antisocial”, se podría decir que sus habilidades de interacción social, comparadas con sus pares, se encuentran poco desarrolladas; otro ejemplo: en vez de decir que un niño tiene una “inteligencia anormal”, se podría referir que su inteligencia cae por debajo de los límites de la norma, o que comparado con los niños de su edad, su desempeño está por debajo de la norma. La segunda recomendación es que no solamente se proporcione el informe o la entrega del reporte de evaluación por escrito, sino que en una sesión programada para tal fin, además de proporcionarle al usuario este reporte, se le expliquen verbalmente los resultados que se encontraron, así como las conclusiones y las sugerencias que se proponen para la toma de decisiones.

De esta manera, si cumplimos con los requisitos profesionales y académicos para llevar a cabo una evaluación psicológica, a saber: estar bien capacitado en el uso e interpretación de pruebas, aplicar sólo las que cumplan con las garantías científicas que la psicometría establece y seguir las normas éticas establecidas para la profesión, el

psicólogo dedicado a la evaluación realizará un ejercicio responsable de su profesión.

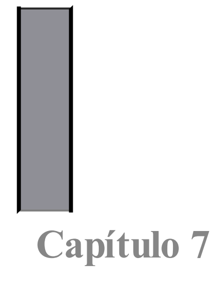
En el siguiente y último capítulo de esta obra, presentamos una guía con recomendaciones y sugerencias para integrar una batería de evaluación psicológica.

Referencias

Aragón, L.; Olivares, R.; Rodríguez, N. y Romero, M. (2002). La entrevista como instrumento de evaluación psicológica. En Aragón, L. E y Silva, A., *Fundamentos teóricos de la evaluación psicológica.* México: Pax

Groth-Marnat, G. (1997). *Handbook of Psychological Assessment.* New York: John Wilely & Sons, Inc.

Sattler, J. M. (1996). *Evaluación infantil.* México: El Manual Moderno.



Capítulo 7

GUÍA PARA SELECCIONAR UNA BATERÍA DE EVALUACIÓN PSICOLÓGICA

Este capítulo tiene como finalidad proporcionar algunas sugerencias y recomendaciones para seleccionar una batería de pruebas cuando se va a llevar a cabo una evaluación psicológica. Una primera fase muy importante de este proceso es la entrevista y la realización de la historia clínica, ya que tenemos que considerar la información obtenida en dicha sesión para poder establecer el motivo de la evaluación. Por ejemplo, algunas de las causas por las que comúnmente se remiten a niños o adolescentes son: bajo rendimiento escolar, problemas de aprovechamiento o aprendizaje, niños inquietos y que no prestan atención, problemas de conducta, mal comportamiento, agresividad, problemas de lenguaje, niños o adolescentes aislados o que no tienen amigos, lectura y escritura deficientes, confusión de letras, tristeza, aflicción, soledad, elección de carrera, entre otros.

El motivo de consulta es aquello que los padres nos plantean como la principal razón por la que acuden a solicitar el servicio de evaluación psicológica; sin embargo, tanto la información recabada como lo observado durante la realización de las entrevistas llevadas a cabo con los padres así como con el niño o el adolescente en cuestión, permiten elaborar hipótesis acerca de que existen otros problemas de conducta que se están presentando además del principal motivo planteado por los progenitores; también podemos preguntarnos qué otros atributos están relacionados con el motivo planteado originalmente o cuáles otros pudieran influir. Por ejemplo, el motivo por el cual pudieran acudir algunos padres a solicitar el servicio es porque su hijo presenta problemas de lectoescritura por confundir letras o grafías y escribir cosas diferentes al tomar dictado. Es claro que a este niño se le tendría que aplicar una prueba para descartar dislexia o bien, para determinar en qué consisten sus errores de tipo disléxico en caso de que los presente; sin embargo, como en la realización de la historia clínica se obtuvieron otros datos referentes a que el niño tiene muy bajas calificaciones, su desarrollo fue un poco demorado y que se muestra muy preocupado por sus dificultades escolares, y además durante la entrevista que le realizamos lo observamos inquieto, preocupado y con conductas como temblores y morderse las uñas. En este caso, además de la prueba de

dislexia, debería aplicarse una de inteligencia para descartar que sus dificultades se deban a bajas capacidades intelectuales, así como para conocer con qué habilidades sí cuenta el niño; aplicaríamos también una prueba de percepción visoespacial para descartar si presenta algunas dificultades en este rubro que pudieran influir en los errores que comete; finalmente, también sería conveniente realizar una prueba de ansiedad para determinar cómo intervienen sus dificultades escolares en su estado emocional y si de hecho la ansiedad pudiera exacerbar todavía más el problema.

En este capítulo no pretendemos plantear todos los tipos de problemas que pudieran presentar ya sea los niños, adolescentes o adultos que acuden a consulta, por lo que para poder ayudar en la selección de una batería de pruebas de evaluación psicológica se describirán a grandes rasgos lo que miden aquellas más utilizadas y que en la actualidad se pueden comprar en México (editadas principalmente por la editorial mexicana El Manual Moderno y por TEA Ediciones de España). De cada prueba vamos a proporcionar su nombre completo, los autores de la misma, la editorial que la publica, el objetivo de ésta, una breve descripción de las subescalas que comprende, qué evalúa cada una de ellas, la información que proporcionan, el tipo de puntuaciones que utilizan y para qué tipo de problemas se recomienda su aplicación.

Por cuestiones didácticas, primero serán presentadas las pruebas del área intelectual, las que miden atributos tales como las capacidades intelectuales, cognitivas o del área académica y en segundo lugar las del área emocional o clínica; describiremos entonces pruebas de inteligencia, aptitudes para el aprendizaje, percepción visoespacial, dislexia, hábitos de estudio y orientación vocacional; y pruebas que evalúan ansiedad, problemas de conducta social y depresión. La información que se presenta en este capítulo se encuentra en los manuales de las pruebas citadas.

Esperamos que esta guía sea de utilidad para seguir formándose en evaluación psicológica, un área tan importante e indispensable para la intervención en dicha disciplina.

Pruebas del área intelectual o académica

Evaluación de la inteligencia

En México se pueden utilizar cuatro pruebas para medir inteligencia: las tres escalas de David Wechsler: WISC-IV, WPPSI-III y WAIS-IV (editadas por la editorial El Manual Moderno) y la Escala de Inteligencia 3 Factor “g” de R. B. Cattell y A. K. S. Cattell (editada por TEA); las tres de Wechsler y la de Cattell tienen normas para México. Empezaremos describiendo la Escala Wechsler de Inteligencia para Niños IV, que es la segunda prueba utilizada con mayor amplitud en todo el mundo y aquí en México no es la excepción.

Wechsler afirma que obtener un CI no es la única finalidad de las pruebas de inteligencia ni es tampoco la más importante; además tiene como objetivo la evaluación de la capacidad del sujeto para comprender el mundo que le rodea y los recursos que

posee para enfrentarse con sus exigencias y sus desafíos. Así, para este autor, la inteligencia es la capacidad global, multideterminada y polifacética de un individuo, tendiente a entender y enfrentar el mundo que le rodea; es decir, es la suma de muchas habilidades, las cuales se han desarrollado de diferentes formas y se pueden expresar de diversas y múltiples maneras. Su finalidad es poder desenvolverse adecuadamente para enfrentar los retos que se presentan día a día en diferentes contextos (escolar, social, vida cotidiana, entre otros).

Podemos considerar, entonces, que la capacidad intelectual de las personas se demuestra al enfrentarse a diferentes situaciones y tareas en tres campos problemáticos específicos:

a) Académicos y técnicos: como el trabajo escolar, la ciencia, las profesiones, las especialidades ocupacionales.
b) Prácticos: se refiere a los problemas que se plantean al afrontar los asuntos cotidianos: ganar dinero, planificar tanto diversas acciones como la propia vida.
c) Sociales: son los que surgen en la interacción con los otros individuos, grupos o incluso culturas extrañas.

De esta manera, ser inteligente significa disponer tanto de recursos para resolver problemas prácticos y académicos como estar capacitado para entender, pensar y relacionarse adecuadamente con los demás.

Wechsler menciona que existen factores no intelectuales que pueden afectar la ejecución y las capacidades; consisten en conglomerados de aptitudes que, más que habilidades cognoscitivas, son rasgos de personalidad, que aunque son independientes de la habilidad intelectual, sirven para dirigir y aumentar la utilización de esas capacidades intelectuales (p. ej., persistencia, gusto, control de impulsos, autocontrol, motivación, entre otros). De acuerdo a la concepción anterior, actuar de modo inteligente implica cierto grado de autodisciplina para ser capaz de pensar antes de actuar, para ser crítico con el propio modo de actuar, para mantener la atención puesta en la meta a conseguir, para persistir en la tarea pese a las dificultades; por eso es que en la mayoría de las ocasiones es indispensable, además de una prueba de inteligencia, aplicar otras relacionadas con aspectos emocionales.

La Escala Wechsler de Inteligencia para el nivel escolar cuarta edición (WISC-IV) es un instrumento clínico de aplicación individual para evaluar la inteligencia en niños de seis años 0 meses a 16 años 11 meses de edad (6:0 – 16:11). Ésta es una revisión de la Escala Wechsler de Inteligencia para el nivel escolar tercera edición (WISC-III, 1991). La estandarización mexicana la coordinó el Dr. Pedro Antonio Sánchez Escobedo. La muestra final quedó integrada por 1 234 participantes, de los cuales 87% de estudiantes provino de escuelas públicas y el resto de planteles privados; la muestra se seleccionó en 12 estados de cinco diferentes regiones del país: sureste (Yucatán y Campeche), occidente (Michoacán, Colima y Jalisco), noroeste (Sonora y Sinaloa), centro-norte (Coahuila, San Luis Potosí y Aguascalientes) y centro (Distrito Federal y Morelos).

WISC IV proporciona puntuaciones compuestas que representan el funcionamiento intelectual en dominios cognoscitivos específicos, estas puntuaciones son típicas y están basadas en la curva normal. El rango de la norma, tanto para esta escala como para las otras dos de Wechsler, es de 90 a 110 (con media de 100), tanto para el CI Total como para los Índices; para cada subprueba, la norma de las puntuaciones estandarizadas es de siete a 13 (con media de 10). La prueba proporciona cinco CI´s o Índices Compuestos:

1. Índice de Comprensión verbal: mide capacidades verbales al utilizar razonamiento, comprensión y formación de conceptos.
2. Índice de Razonamiento perceptual: mide razonamiento y organización perceptuales.
3. Índice de Memoria de trabajo: evalúa atención, concentración y memoria.
4. Índice de Velocidad de procesamiento: mide velocidad de procesamiento mental y grafomotor.
5. Índice o capacidad intelectual general o CI Total.

A continuación se muestra una tabla en donde se presentan las cinco puntuaciones Índices o Escalas que proporciona la prueba, así como las subpruebas que las conforman; las que están entre paréntesis y con cursivas son pruebas suplementarias.

Tabla 7-1. Índices y subpruebas de WISC-IV

Índice o Escala	**Subpruebas**
COMPRENSIÓN VERBAL (*ICV*)	Semejanzas (*SE*)
	Vocabulario (*VB*)
	Comprensión (*CM*)
	(Información) (*IN*)
	(Palabras en contexto [Pistas]) (*PC*)
RAZONAMIENTO PERCEPTUAL (*IRP*)	Diseño con cubos (*DC*)
	Conceptos con dibujos (*CD*)
	Matrices (*MT*)
	(Figuras incompletas)(*FI*)
MEMORIA DE TRABAJO (*IMT*)	Retención de dígitos (*RD*)
	Sucesión de números y letras (*NL*)
	(Aritmética) (*AR*)
VELOCIDAD DE PROCESAMIENTO (*IVP*)	Claves (*CL*)
	Búsqueda de símbolos (*BS*)
	(Registros) (*RG*)
ESCALA TOTAL (*CI*)	Diez pruebas esenciales: DC, SE, RD, CD, CL, VB, NL, MT, CM y BS

WISC-IV es una escala de inteligencia que proporciona mucha información; como ya mencionábamos anteriormente, no sólo importa el CI total y los cuatro Índices compuestos, esta prueba no se aplica solamente para conocer la cantidad del atributo que un niño o adolescente posee, sino que deben interesarnos los procesos cognitivos que el niño tiene y de los que carece; por ello, enseguida proporcionamos de manera resumida

la información que se obtiene de cada una de las subpruebas de las cuales consta el instrumento.

Subpruebas de comprensión verbal

Semejanzas. Mide razonamiento verbal y formación de conceptos; también implica comprensión auditiva, memoria, distinción entre características esenciales, y no esenciales y expresión verbal.

Vocabulario. Mide conocimiento de palabras y formación de conceptos verbales; reserva de conocimientos, capacidad de aprendizaje, memoria a largo plazo, grado de desarrollo del lenguaje.

Comprensión. Mide razonamiento y conceptualización verbales, comprensión y expresión verbales, capacidad para evaluar y utilizar la experiencia pasada y para demostrar información práctica. Implica conocimiento de las normas sociales de comportamiento, madurez, juicio social y sentido común.

Información. Mide la capacidad para adquirir, conservar y recuperar conocimiento general, escolar y del ambiente. Implica inteligencia cristalizada, memoria a largo plazo, así como capacidad de expresión verbal.

Palabras en contexto (Pistas). Mide razonamiento verbal, comprensión verbal, capacidad analógica y de razonamiento general, abstracción verbal, capacidad para integrar y sintetizar diferentes tipos de información y para generar conceptos alternativos.

Subpruebas de razonamiento perceptual

Diseño con cubos. Mide capacidad de análisis y síntesis de estímulos visuales abstractos. Implica formación de conceptos no verbales, percepción y organización visuales, procesamiento simultáneo, coordinación visomotora, capacidad de separar figura y fondo en los estímulos visuales. Implica observación visual así como la capacidad de integrar procesos visuales y motores.

Conceptos con dibujos. Mide capacidad de razonamiento abstracto y categórico.

Matrices. El tipo de tareas de analogías de matrices que presenta la subprueba constituyen buenas medidas de la inteligencia fluida y estimados confiables de la capacidad intelectual general; están relativamente libres de influencias socioculturales y del lenguaje, y no requieren manipulación.

Figuras incompletas. Mide percepción y organización visual, concentración, sentido común y reconocimiento de los detalles esenciales de los objetos.

Subpruebas de memoria de trabajo

Retención de dígitos. Mide memoria auditiva a corto plazo, habilidades de

secuenciación, atención y concentración, transformación de la información, manipulación mental y formación de imágenes visoespaciales.

Sucesión de números y letras. La tarea implica manipulación mental, atención, memoria auditiva a corto plazo, formación de imágenes visoespaciales y velocidad de procesamiento.

Aritmética. Implica manipulación mental, concentración, atención, memoria a corto y largo plazo, capacidad de razonamiento numérico y atención mental.

Subpruebas de velocidad de procesamiento

Claves. Mide memoria a corto plazo, capacidad de aprendizaje, percepción visual, coordinación visomotora, capacidad de rastreo visual, flexibilidad cognitiva, atención y motivación.

Búsqueda de símbolos. Mide memoria visual y a corto plazo, coordinación visomotora, flexibilidad cognoscitiva, discriminación visual y concentración.

Registros. Mide atención visual selectiva, vigilancia y descuido visual.

Se recomienda aplicar esta prueba y en general una de inteligencia, cuando el motivo de consulta sea por problemas o dificultades escolares: niños o adolescentes que no aprenden o tienen problemas de aprendizaje, que tienen muy bajas calificaciones o bajo rendimiento escolar, cuando hay sospecha de bajas capacidades intelectuales y/o su desarrollo ha sido demorado. Igualmente si hay sospecha de que un niño tiene una inteligencia sobredotada o superior.

WPPSI-III es una escala de inteligencia para niños preescolares en edades comprendidas entre los dos años seis meses hasta los siete años tres meses, y es una revisión de la Escala WPPSI-R de 1989. En México, la estandarización la llevó a cabo la Dra. Carolina Uribe Ferrari, con una muestra estratificada de 829 niños de escuelas públicas y privadas provenientes de 21 estados de la República Mexicana: Baja California Norte, Sonora, Sinaloa, Coahuila, Chihuahua, Durango, San Luis Potosí, Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, DF, Morelos, Puebla, México, Chiapas, Guerrero, Oaxaca, Tabasco y Veracruz.

Esta escala tiene dos niveles de aplicación: para niños de dos años seis meses a tres años 11 meses y la otra para niños de cuatro años a siete años tres meses. La primera se recomienda aplicar sólo si se sospecha que el niño tiene capacidades cognitivas o intelectuales superiores. La segunda comúnmente se aplica a niños de cuatro a seis años, ya que se superpone con la Escala WISC-IV, pero se recomienda aplicar WIPPSI-III a niños mayores de seis años hasta los siete años tres meses si se sospecha que el niño de esas edades tiene limitadas capacidades intelectuales o bien, si cuando al empezar a aplicar WISC-IV, no se obtienen puntuaciones acertadas. WPPSI-III brinda puntuaciones de un CI Verbal, un CI de Ejecución y un CI Total y, para niños de cuatro a siete años tres meses, un Índice de Velocidad de Procesamiento.

Para los niños de 2:6 a 3:11 años, la Escala Verbal se compone de las subpruebas

Vocabulario receptivo e Información; para la Escala de Ejecución las subpruebas son Diseño con cubos y Rompecabezas; existe una prueba suplementaria de la Escala Verbal que es Denominaciones.

Para los niños de cuatro años a 7:3 meses, la Escala Verbal se compone de las subpruebas: Información, Vocabulario y Pistas, con dos pruebas suplementarias: Comprensión y Semejanzas. La Escala de Ejecución la componen las subpruebas Diseño con cubos, Matrices y Conceptos con dibujos, y tiene dos pruebas suplementarias: Figuras incompletas y Rompecabezas. El Índice de Velocidad de Procesamiento se compone de las pruebas Claves y la suplementaria Búsqueda de símbolos. Enseguida se describirá lo que evalúan las subpruebas Vocabulario receptivo, Denominaciones y Rompecabezas; lo que evalúan las demás subpruebas es lo mismo que lo expuesto en WISC-IV.

Vocabulario receptivo evalúa la habilidad para comprender las instrucciones verbales, la discriminación auditiva y visual, la memoria auditiva y el procesamiento auditivo y auditivo-visual.

Denominaciones evalúa el lenguaje expresivo, la recuperación de palabras de memoria a largo plazo y la asociación de estímulos visuales con el lenguaje.

Rompecabezas mide la organización perceptivo visual, la integración y la síntesis de las relaciones de la parte con el todo, razonamiento no verbal, aprendizaje por ensayo y error, coordinación visomotora, capacidad de planeación y persistencia en la tarea.

WAIS-IV es una escala de inteligencia para adultos; se puede aplicar desde los 16 a los 90 años de edad. Tiene los mismos índices que WISC-IV: Comprensión Verbal (medido a través de las subpruebas Vocabulario, Semejanzas, Información y Comprensión); Razonamiento Perceptual (que comprende las subpruebas Figuras incompletas, Peso figurado, Diseño con cubos, Matrices y Ensamble de objetos); Memoria de Trabajo (medido a través de de Retención de dígitos, Aritmética y Sucesión de números y letras) y Velocidad de Procesamiento (medido a través de Búsqueda de de símbolos, Claves y Cancelación). Al igual que en WIPPSI, la descripción de lo que evalúa cada subescala se puede consultar en WISC-IV.

Al igual que con WISC-IV y WPPSI-III, se recomienda aplicar esta prueba a adolescentes de 16 años en adelante que presenten dificultades escolares. En adultos, esta prueba se emplea en la selección de personal, para determinar sus capacidades intelectuales, y en algunas ocasiones para medir el deterioro cognoscitivo que los adultos pueden presentar al envejecer, y en qué áreas sucede.

La Escala de Inteligencia 3 Factor "g" de R. B. Cattell y A. K. S. Cattell es una prueba no verbal que mide la inteligencia fluida, es decir, la inteligencia libre de influencias socioculturales. En las bases para la construcción de esta prueba, estaba originalmente la necesidad de un instrumento que midiera con precisión la inteligencia de sujetos con diferente lenguaje, cultura y nivel social o educativo; para Cattell y Cattell, la reducción de los efectos debidos a estas diferencias era una meta deseable al evaluar la inteligencia, de hecho, el nombre original de esta prueba es *Culture Fair Intelligence Tests,* esto es, Prueba de Inteligencia libre de influencias culturales.

De esta manera, en la construcción de su prueba, Cattell y Cattell intentaban buscar una medida más pura de inteligencia, un instrumento que les permitiera eliminar los efectos culturales o ligados a la información para mantener los más puramente procesuales; esto aportó como resultado aquellas aptitudes vinculadas al proceso lógico, a la inducción y a la generalización, minimizó factores ligados a la memoria y a aspectos socioculturales, para poder establecer una separación más clara entre la aptitud natural y el aprendizaje específico, de este modo fue posible realizar un mejor análisis del potencial individual.

La investigación original respecto al desarrollo de esta prueba, muestra que aquéllas que están constituidas por elementos gráficos apenas son alteradas por influencias culturales, en tanto que en las pruebas verbales o manipulativas es difícil eludir la influencia cultural.

Teóricamente, el factor "g" o factor general de inteligencia, está implicado en una gran variedad de tareas cognitivas, tales como descubrir semejanzas y diferencias entre elementos, inferir correspondencias, reglas y generalizaciones, así como seguir una línea de razonamiento y predecir consecuencias. Por otra parte, dicho factor sugiere que la naturaleza de las diferencias individuales puede ser entendida en términos de los niveles de energía mental que el individuo utiliza en la realización de una tarea intelectual y/o en términos de las aptitudes que el sujeto utiliza para realizar la tarea a partir de tres principios cualitativos: percepción de la experiencia, deducción de relaciones y de elementos correlacionados.

En este contexto es que se encuentran los trabajos realizados por Cattell y Cattell, quienes analizan el problema de la medición desde lo que denominan **inteligencia fluida** (capacidad natural o básica) e **inteligencia cristalizada** (capacidades adquiridas a través del aprendizaje y la experiencia). La primera implica la posesión por parte del sujeto, de una **cantidad suficiente de inteligencia general** que le va a permitir adquirir los niveles de aptitud suficientes en lenguaje, matemáticas y otros sistemas simbólicos, posible sólo si se da la suficiente exposición de éstos a través de la educación, el entrenamiento y la experiencia.

A su vez y de acuerdo con el grado en que el sujeto se haya beneficiado de esta exposición, generará la llamada **inteligencia cristalizada.** Finalmente, el grado en que una persona se enfrente a diversas tareas y problemas, reflejará sus aptitudes generales tanto de inteligencia fluida como cristalizada.

Como ya mencionamos al inicio de este apartado, esta prueba tiene normas para la población mexicana, tanto de bachillerato como de licenciatura. La baremación fue realizada por las Mtras. Laura Edna Aragón Borja y Ma. del Rosario González Velázquez. La muestra del estudio la conformaron alumnos de bachillerato y licenciatura, provenientes de 34 diferentes escuelas públicas de la zona metropolitana de la Cd. de México.

La muestra del bachillerato se realizó con 895 alumnos de los tres grados escolares de 22 escuelas, 12 del turno matutino y 10 del vespertino; seis escuelas de la UNAM: Escuelas Nacionales Preparatorias cuatro, cinco y nueve, y Colegios de Ciencias y

Humanidades Naucalpan, Sur y Oriente; tres Colegios de Bachilleres: uno, dos y 12; tres Centros de Estudios Científicos y Tecnológicos (CECyT) del Instituto Politécnico Nacional: dos, ocho y 15; dos Preparatorias Oficiales del Estado de México: Atizapán y Tultitlán; dos Centros de Estudios Tecnológicos, Industriales y de Servicios (CETIS 5 y 95); dos Centros de Bachillerato de Estudios Tecnológicos, Industriales y de Servicios (CEBETIS 160 y 202) y cuatro Colegios Nacionales de Estudios Profesionales: CONALEPs Gustavo A. Madero, Atizapán, Netzahualcóyotl III y México-Canadá.

La muestra de licenciatura la integraron 427 alumnos de 12 escuelas: Tecnológico Regional de Tlalnepantla, Escuela Superior de Turismo del IPN, Escuela Superior de Ingeniería del IPN, Universidad Autónoma Metropolitana Campus Azcapotzalco, Iztapalapa y Xochimilco, Universidad Autónoma del Estado de México y las cinco Facultades de Estudios Superiores de la UNAM: Acatlán, Aragón, Cuautitlán, Iztacala y Zaragoza. En dichas escuelas, se aplicó la prueba en las siguientes licenciaturas: Administración, Turismo, Ingeniería Química, Arquitectura, Diseño industrial, Diseño gráfico, Ciencias biológicas, Ciencias sociales, Contaduría, Ingeniería mecánica, Ingeniería eléctrica, Pedagogía, Actuaría, Optometría, Economía y Relaciones Internacionales.

La Escala de Inteligencia 3 Factor “g” de Cattell y Cattell es una prueba muy sencilla y rápida de aplicar, toma aproximadamente media hora, además de que su aplicación puede ser tanto individual como colectiva. Consta de cuatro subpruebas:

1. Series. La tarea del evaluado consiste en seleccionar entre las opciones que se le presentan, el dibujo que continúa una serie; su duración es de tres minutos.
2. Clasificación. El sujeto tiene que escoger, de una serie de cinco figuras, aquellas dos que no concuerdan o son diferentes con respecto a las otras tres; ésta es una tarea de formación de conceptos y su duración es de cuatro minutos.
3. Matrices. La tarea consiste en completar la matriz incompleta de dibujos que se le presentan al sujeto, mediante la elección de una de las cinco opciones que se le proponen; dura tres minutos.
4. Condiciones. Aquí el evaluado tiene que seleccionar una alternativa de las que se le presentan, en donde se podría cumplir la misma condición que se muestra en la figura estímulo; su duración es de 2.5 minutos.

Esta prueba proporciona puntuaciones estandarizadas basadas en una curva normal, un CI con media de 100 y desviación típica de 15; aquí las puntuaciones normativas son de 85 a 115; además, brinda puntuaciones percentiles, es decir, con la puntuación obtenida por cada sujeto evaluado, es posible determinar tanto su CI y si éste se encuentra en la norma, o bien, por debajo o por arriba de ella, y con el rango percentil se puede saber qué tan abajo o arriba de la norma se encuentra; esto es, el porcentaje de sujetos de su población que se encuentran por arriba de la persona evaluada, o bien, al que sobrepasa.

Los autores recomiendan aplicar esta prueba a estudiantes de preparatoria o nivel superior, para averiguar si las realizaciones del sujeto son las que cabe esperar de su

inteligencia; por tanto, facilita la identificación de posibles problemas emocionales o de aprendizaje. Además permite determinar el potencial de candidatos para realizar tareas en las que esté implicada la aptitud cognitiva y posibilita la selección de aspirantes en programas de becas o de instrucción superior. Como es de fácil aplicación, de un tiempo relativamente corto y además se puede aplicar colectivamente, también se sugiere aplicarla cuando por cuestiones de tiempo no sea posible utilizar WISC-IV o WAIS-IV, para selección de personal, como parte de una batería de orientación vocacional, así como cuando es necesario evaluar grandes grupos de sujetos; por ejemplo, en la UEPI la utilizamos como parte de una batería para el ingreso de aspirantes a especialidades odontológicas y también la hemos aplicado con aspirantes a las residencias médicas.

Evaluación de las aptitudes para el aprendizaje

Podemos considerar que las aptitudes para el aprendizaje son aquellas habilidades precurrentes o necesarias para que los conocimientos que se enseñan en la escuela puedan ser adquiridos por los niños, esto es, para que el aprendizaje escolarizado tome lugar. En este apartado revisaremos dos pruebas: la Batería de Aptitudes para el Aprendizaje Escolar Nivel 1, de Ma. Victoria de la Cruz (BAPAE, editada por TEA y con normas para México), y la Prueba de Percepción visoespacial de Frostig-II (DTVP-2, editada por la editorial El Manual Moderno).

Batería de aptitudes para el aprendizaje escolar (BAPAE)

Esta prueba evalúa el repertorio de habilidades académicas básicas con que cuenta un niño, con la finalidad de poder determinar si son suficientes para que éste afronte las demandas del aprendizaje de la lectoescritura y aritmética básica, por lo que esta prueba se aplica a niños de seis o siete años que se encuentran en el último año del Jardín de Niños, o bien, al inicio del primer año escolar.

Esta prueba tiene validez predictiva, es decir, intenta predecir si el niño tendrá éxito en el primer año de la enseñanza escolarizada y, en caso de obtener puntuaciones bajas, el objetivo será subsanar las deficiencias encontradas para prevenir que el niño, al no aprender a leer y escribir de manera eficiente, presente a corto y largo plazo bajo rendimiento académico y fracaso escolar.

La muestra para obtener las normas para México la conformaron 1 209 niños, 643 niños (53.2%) y 566 niñas (46.8%), de seis a siete años de edad que cursaban el primer año de Educación Primaria, los cuales fueron seleccionados entre 45 escuelas públicas elegidas aleatoriamente de 11 de las 16 delegaciones políticas del Distrito Federal y de cuatro municipios del Estado de México que conforman la zona metropolitana.

BAPAE es una prueba no verbal, en la que se utiliza un cuadernillo con dibujos en donde el niño tiene que marcar la o las figuras que se le mencionan; se puede administrar en forma tanto individual como colectiva, en aproximadamente una hora, y consta de tres grandes áreas:

Comprensión verbal. Evalúa la extensión del vocabulario receptivo del niño así como conceptos.

Aptitud numérica. Evalúa conceptos cuantitativos (tamaño, lugar, cantidad) y manejo de números (pequeñas sumas y restas).

Aptitud perceptiva y espacial. A su vez comprende Relaciones espaciales (el niño debe elegir una figura para que al ser colocada en el espacio y la posición correctos complete una imagen para igualarla al dibujo muestra), Constancia de forma (elegir los dibujos exactamente iguales al de la muestra) y Orientación espacial (elegir los dibujos que estén colocados en la misma posición que el dibujo muestra).

Esta prueba proporciona cuatro puntuaciones en percentiles por cada área, a saber: Comprensión verbal, Aptitud numérica, Aptitud perceptiva y una Puntuación total que es la suma de las tres aptitudes anteriores. Con ellas podemos determinar cuáles son las deficiencias que ese niño tiene y entonces qué ejercicios habría que programar para prevenir que éste no aprenda a leer o escribir adecuadamente, en caso de que obtenga puntuaciones por debajo de la norma (menores a 25), o bien, se encuentre en la norma baja (de 26 a 35).

La autora de la prueba menciona que es importante considerar que los resultados de esta prueba no deben ser utilizados para establecer una calificación inamovible de los niños, ya que las aptitudes que se aprecian con ésta tienen carácter de adquiridas, lo cual significa que son susceptibles de ser modificadas; la primordial intención es localizar a niños con atrasos en estas áreas para ayudarles a alcanzar el nivel normal mediante actividades o programas de desarrollo en estos aspectos.

Es recomendable aplicar esta prueba con fines preventivos al final del Jardín de Niños, para asegurarnos que el niño no tenga problemas en el aprendizaje de los contenidos del primer año de enseñanza básica. Por otro lado, también se sugiere cuando un niño presenta dificultades en la escuela y se ha descartado ya, a través de una prueba de inteligencia, que tiene bajas capacidades intelectuales; porque entonces cabría suponer que quizás no ha desarrollado esas capacidades al máximo. Asimismo, cuando un niño presenta dificultades para aprender, se aplica BAPAE y, si sus puntajes son bajos, posteriormente se puede realizar una prueba de inteligencia para complementar la información.

Método de evaluación de la percepción visual (DTVP-2)

En los seres humanos –y en la mayoría de los animales–, los órganos de los sentidos se especializan en la recepción de tipos específicos de estímulos: nuestra piel es sensible al calor, al frío, a la presión y a la luz; tenemos lengua para gustar, nariz para oler, oídos para escuchar, ojos para ver; sin embargo, la percepción visual es más que sólo ver: percibir es el acto de **interpretación** de un estímulo, recibido por el cerebro a través de uno o más de los órganos de los sentidos, por lo tanto, el término percepción visual no significa ver con precisión, puesto que la interpretación de los estímulos visuales ocurre en el cerebro y no en los ojos; en consecuencia, lo que percibimos en un momento dado,

depende no sólo de la naturaleza del estímulo existente, sino también de los antecedentes que le afectan: nuestras experiencias pasadas, nuestros sentimientos del momento, nuestros prejuicios, deseos y actitudes.

La importancia de la percepción reside en que las valiosas experiencias acumuladas a través del desarrollo, van a propiciar que el individuo pueda adquirir conocimientos nuevos, los cuales le van a permitir dar la respuesta adecuada ante situaciones o problemas novedosos. Marianne Frostig considera a la percepción visual como la capacidad de reconocer y discriminar estímulos visuales, y de interpretarlos cuando se asocian con experiencias previas.

El período que va del nacimiento a los dos años aproximadamente, es el de máximo desarrollo motor; entre el año y medio a los tres años y medio se produce la mayor evolución del lenguaje, y entre los tres años a los ocho años y medio, se da el máximo desarrollo perceptual, siendo este lapso en el que se inicia la etapa escolarizada del niño.

En la Teoría del Procesamiento de la Información (en la que está basada esta prueba), la percepción es el paso intermedio entre la sensación y la cognición. En la primera, el individuo únicamente debe tener intactos los órganos de los sentidos. En esta fase detectamos la presencia de un estímulo mediante la vista, el oído, el tacto, el olfato, el gusto; únicamente se atiende a los aspectos sensoriales. Sin embargo, en la percepción no sólo detectamos la presencia, sino que además interpretamos y organizamos los elementos físicos de los estímulos que hemos detectado, sus propiedades: tamaño, color, forma, textura, sonido, entre otros; aquí intervienen operaciones cerebrales que involucran la interpretación y organización de los estímulos. Finalmente, en la cognición, se atiende a los aspectos simbólicos de los estímulos detectados y percibidos, involucran pensamiento, lenguaje significativo o solución de problemas.

Para los autores de la prueba, el término percepción se refiere a las operaciones cerebrales que involucran la interpretación y organización de los elementos físicos de un estímulo, y que pueden diferenciarse fácilmente de la sensación (como la agudeza visual y la sensibilidad auditiva) y de las tareas de procesamiento cognitivo (atención, comprensión de lectura).

Por eso la percepción visual es un repertorio básico para el aprendizaje de la lectoescritura: el niño tiene, en primer lugar, que ver correctamente las letras, discriminarlas y distinguirlas de otro tipo de letras o estímulos visuales, posteriormente deberá interpretarlas para finalmente poder asignarles un significado; es decir, la percepción visual es una precurrente para el aprendizaje de la lectoescritura.

Esta prueba fue elaborada originalmente por Marianne Frostig, y ésta es una versión ampliada y mejorada de la misma llevada a cabo por Donald D. Hammill, Nils A. Pearson y Judith K. Voress, después de la muerte de la autora. La DTVP-2 es una batería de ocho subpruebas, se aplica aproximadamente en una hora y su finalidad es evaluar el estado perceptivo visual de niños de cuatro a diez años de edad a través de dos grandes escalas: la de Habilidades Visomotoras y la de Percepción Visual con Respuesta Motora Reducida. En la primera, el niño resuelve los ejercicios que se le presentan mediante una respuesta grafomotora, mientras que en la segunda, sólo tiene que señalar o

decir cuál es la figura correcta. La muestra normativa constó de 1 972 niños residentes de 12 diferentes estados de la Unión Americana.

A continuación presentamos una tabla en donde se presentan las ocho subpruebas del DTVP-2 y a qué escala pertenecen; el número se refiere al orden de aplicación:

Tabla 7-2. Escalas y subpruebas de DTVP-2

HABILIDADES VISOMOTORAS	PERCEPCIÓN VISUAL CON RESPUESTA MOTORA REDUCIDA
1. Coordinación ojo-mano	2. Posición en el espacio
3. Copia	4. Figura-fondo
5. Relaciones espaciales	6. Cierre visual
7. Velocidad visomotora	8. Constancia de forma

Coordinación ojo-mano mide la habilidad para dibujar líneas rectas o curvas con precisión de acuerdo con límites visuales impuestos.

Copia mide la habilidad de reconocer los rasgos de un diseño y dibujarlo a partir de un modelo.

Relaciones espaciales mide la habilidad para juntar puntos, para reproducir patrones presentados visualmente al unir puntos y respetar el espacio que se les muestra.

Velocidad visomotora mide la rapidez con que un niño puede hacer determinadas señales en ciertos diseños (esta prueba es similar a la de Claves de WISC-IV).

Posición en el espacio mide la habilidad para igualar dos figuras de acuerdo con sus rasgos comunes y seleccionar la figura exacta de una serie de figuras similares, aunque diferentes.

Figura-fondo mide la habilidad de ver figuras específicas cuando están ocultas por un fondo confuso y complejo.

Cierre visual mide la habilidad para reconocer una figura estímulo que ha sido dibujada de manera incompleta; los niños tienen que llenar mentalmente lo que falta en la figura.

Constancia de forma mide la habilidad para igualar dos figuras que varían en uno o más rasgos distintivos (tamaño, posición, sombreado).

DTVP-2 es una de las pruebas que brinda mucha información, proporciona puntuaciones estandarizadas basadas en la curva normal: tres coeficientes o compuestos con una media de 100 y desviación típica de 15: Cociente de Percepción Visual General (PVG), Cociente de Percepción Visual con Respuesta Motora Reducida (PMR) y Cociente de Integración Visomotora (IVM). El primero determina de manera resumida la percepción visual del niño evaluado: si está en la norma (85 a 115), por abajo o por arriba de ella. Los otros dos brindan puntuaciones separadas de las habilidades de percepción del niño: cuando tiene una buena percepción siempre y cuando no esté implicada una respuesta motora (PMR) y la percepción visual en donde sí la integra (IVM). Este dato es interesante porque un niño pudiera tener un bajo PVG pero sólo una de las dos escalas estar baja y la otra en la norma; por ejemplo, podría tener un IVM bajo pero un PMR en la norma, lo cual nos hablaría de que las deficiencias del niño se

presentan en la coordinación visomotriz; igualmente, otro niño pudiera tener un adecuado PVG y, sin embargo, en alguno de los dos compuestos estar bajo, en cuyo caso uno compensaría al otro; en dicho caso habría que trabajar para subsanar las deficiencias en el que está bajo.

Asimismo, en cada uno de los compuestos la prueba proporciona puntuaciones percentiles para determinar en qué porcentaje supera el niño o es superado por la población de referencia, así como la mediana de la edad promedio de ejecución; es decir, si un niño tiene una puntuación baja, nos dice qué tan baja es comparada con los de su edad y, en términos comparativos, a qué edad equivaldría la ejecución que presenta.

Por otro lado, para cada una de las subpruebas proporciona puntuaciones estandarizadas, percentiles y equivalentes de edad; las puntuaciones estandarizadas normativas son de ocho a 12, con media de 10; en cuanto a los percentiles, ya sabemos que la norma es de 26 a 74 y la edad equivalente nos diría la edad promedio de ejecución en la habilidad medida. De esta manera, aparte de saber si un niño tiene una puntuación por debajo de la norma o se encuentra en la norma baja en algunos de los compuestos (PVG, PMR o IVM), podríamos determinar específicamente en qué habilidades se encuentran esas deficiencias, con miras a establecer ejercicios para desarrollarlas.

Es indispensable aplicar esta prueba cuando el niño presenta errores de tipo disléxico, ya que es necesario saber si las fallas que comete son secundarias a una inadecuada percepción visoespacial, en cuyo caso primero habría que trabajar sobre los repertorios que la componen, antes de empezar a corregir aquéllos; si el niño tuviera puntuaciones adecuadas en esta prueba, también sería importante conocerlo porque entonces se podría saber que no tiene problemas de percepción visual y entonces habría que reenseñar la lectoescritura. Igualmente, si un niño tuviera problemas en el aprendizaje de ésta, se podría aplicar esta prueba para determinar si ese problema se debe a una falla en sus repertorios perceptivo-visuales. Finalmente, cuando a un niño, después de aplicarle las pruebas de Wechsler WISC-VI o WPPSI-III, se le detecta que tiene serios problemas ya sea en Razonamiento Perceptual o en la Escala de Ejecución, respectivamente, o bien, en BAPAE en el área de Aptitud Perceptiva y Espacial, es recomendable aplicarle DTVP-2 para tener datos más específicos acerca de los problemas de percepción que pudiese presentar.

Evaluación de la dislexia

La finalidad del lenguaje hablado es la comunicación, sin embargo, materializar el lenguaje hablado a través de la escritura es lo que ha permitido a la humanidad la transmisión de conocimientos de una manera masiva, ahí radica precisamente la importancia de la lectoescritura, ya que es el medio por excelencia utilizado para el aprendizaje escolarizado. Ésta, además, es un proceso que consiste en darle significado a los símbolos gráficos que representan al lenguaje hablado; en la escritura se representa al lenguaje por medio de símbolos gráficos y en la lectura se decodifican esos símbolos y se les dan significado.

Para que un niño pueda aprender a leer y escribir necesita poseer madurez tanto biológica como psicológica y un medio que le proporcione la enseñanza de la lectoescritura. A nivel biológico, un niño necesita contar con un sistema nervioso central intacto (esto es, que no tenga daño), completo (que no le falten partes) y maduro (adecuadamente desarrollado). También debe tener un buen funcionamiento de los órganos de los sentidos (vista y oído), o algún dispositivo correctivo (lentes, aparato auditivo, lectura en Braille, si fuese el caso), así como poseer partes del cuerpo que le permitan escribir.

A nivel psicológico, un niño necesita tener las habilidades intelectuales y cognitivas adecuadas a su edad, lo cual a su vez es precurrente para un óptimo desarrollo del lenguaje (comprensión, expresión y adecuada articulación), necesita realizar discriminaciones finas tanto visuales como auditivas, poseer habilidades de comprensión visual y auditiva, así como de expresión motora, haber desarrollado la coordinación visomotora y auditivo-motora, tener bien establecido el sentido de la dirección o lateralidad, así como conciencia del esquema corporal (ubicación espaciotemporal).

Finalmente, para que un niño aprenda a leer y escribir además de presentar una madurez biológica y psicológica, necesita contar con un medio social que le proporcione la enseñanza de la lectoescritura, este medio por lo general es la escuela.

Cuando hay alteraciones en este proceso de aprendizaje surgen los problemas o dificultades específicos de aprendizaje escolar, uno de ellos es la dislexia. Tradicionalmente se le ha considerado como un desorden manifestado por dificultad para leer, a pesar de instrucción convencional, inteligencia normal, adecuadas oportunidades socioculturales y correcto desarrollo personal socio-afectivo; representa un defecto específico de la maduración tradicionalmente asociado a un daño neurológico; esto es, se infiere algún tipo de daño en el sistema nervioso central.

Sin embargo, desde una aproximación cognitivo conductual, se consideraría que los niños llamados disléxicos exhiben deficiencias al responder a la estimulación lingüístico-impresa apropiada para su edad, no porque no tengan la capacidad de leer y escribir correctamente, sino porque han aprendido de manera deficiente o inapropiada a leer y escribir o bien, las condiciones medioambientales han sido las inadecuadas; esto es, desde esta aproximación, la dislexia no es una entidad con existencia propia, no es un problema congénito sino aprendido, y por tanto desde esta perspectiva, no habría niños disléxicos, sino aquellos que cometen errores de tipo disléxico; los disléxicos no son los pequeños, sino los errores que éstos cometen. De esta manera, estos chicos han aprendido a leer y escribir de manera deficiente, cometiendo errores en su lectoescritura que consisten en confundir unos por otros algunos grafemas y fonemas, quizá debido a estrategias de enseñanza inadecuadas, insuficientes ejercicios de práctica para dominar lo enseñado, repertorios discriminativos deficientes, entre otros.

Una ventaja de conceptualizar a la dislexia de esta forma, es que entonces se consideraría que los déficit no necesariamente son permanentes, sino que pueden ser corregidos y modificados.

El Instrumento para Detectar Errores de Tipo Disléxico (IDETID-LEA), de Laura

Edna Aragón Borja, es una prueba orientada al criterio, esto es, no es normativa; el criterio es que el niño lea y escriba correctamente, sin cometer errores de tipo disléxico. Es una prueba dirigida específicamente a la población de alumnos mexicanos de 2° a 5° grados de primaria, donde se evalúan repertorios terminales; es decir, la de 2° año se aplica al terminar 1° y así sucesivamente. La prueba fue validada (validación cruzada) con una muestra total de 660 niños en dos aplicaciones: la primera se realizó a 300 estudiantes de 2º, 3º y 4º grados, con un promedio mínimo de 8 en al área de Español, de nueve diferentes escuelas públicas del D.F., 100 niños por grado escolar y, dentro de cada grado, a 20 se les aplicó copia, a 40 lectura y a 40 dictado. La segunda aplicación se realizó con 360 niños de 21 diferentes escuelas del DF. y área metropolitana, a 120 de ellos se les evaluó copia, a 120 lectura y a 120 dictado; de esos 120 niños, 30 por cada grado escolar, de 2º a 4º y, de cada grupo de 30 alumnos, 15 con promedio de 8, 10 con 9 y, 5 con promedio de 10 en el área de Español.

La prueba consta de tres grandes áreas de contenido: Lectura, Copia y Dictado. A su vez, cada una de ellas consta de cinco Universos de generalización: sílabas, palabras, enunciados, prosas y versos (en Dictado, en lugar de sílabas, son vocales). Por otro lado, la prueba tiene muy bien delimitados y definidos operacionalmente los errores de tipo disléxico que se pueden presentar en cada una de las áreas de contenido; los Universos de generalización están construidos de tal modo, que el niño tiene la oportunidad de presentar todos los tipos de errores disléxicos que se pueden cometer.

Los errores en Lectura pueden ser: Omisiones, Inserciones, Errores de Secuenciación, Confusión por la posición o dirección de los elementos que formas las letras, Confusión por el tamaño de las letras, Confusión por el número de elementos que forman las letras, Separación inadecuada y Confusiones diferentes. Los errores en Copia son los mismos que en Lectura, a lo que se agrega un error: Confusión de mayúsculas y minúsculas. Los errores considerados en Dictado son: Omisiones, Inserciones, Errores de Secuenciación, Confusión de letras minúsculas y mayúsculas, Separación inadecuada de letras y sílabas dentro de la palabra o de los rasgos de las letras, Trazo inadecuado de letras, Confusión por pronunciación similar y Confusiones diferentes.

El IDETID-LEA se aplica idealmente en seis sesiones, aunque también es factible aplicarlo en cuatro, pero con descansos para el niño. Una vez aplicado y calificado, los datos son codificados en tres matrices de resultados incorrectos, una para cada área de contenido, como la que aparece a continuación.

Tabla 7-3. Matriz de resultados incorrectos en lectura

		UNIVERSOS DE GENERALIZACIÓN				
		Sílabas	**Palabras**	**Enunciados**	**Prosas**	**Versos**
Indicadores de error	Omisiones					
	Inserciones					
	Secuenciación					
	Dirección					
	Tamaño					
	No. elementos					
	Separación					

Diferentes					

En cada celda de la matriz se anota el tipo de error que el niño presenta, en qué Universo se presenta y en qué consistió específicamente el error. Una vez que las tres matrices están llenas, se puede proceder a hacer un análisis, tanto cualitativo como cuantitativo, de los tipos de errores que el niño comete, con miras a plantear objetivos de intervención.

Esta prueba se aplica siempre que un niño presenta, ya sea en su lectura o su escritura, errores que consisten en omitir, insertar o cambiar letras, sílabas o palabras por otras. No da un diagnóstico de dislexia puesto que ése no es su objetivo, sino establece si el niño comete errores de tipo disléxico, en qué consisten, y en qué universos de generalización y áreas de contenido se presentan. Esta prueba está publicada en el libro *Evaluación y tratamiento de niños disléxicos* (Editorial Trillas). Se sugiere además aplicar primero el DTVP-2, como ya se sugirió anteriormente al citar dicha prueba y, en caso de considerarse necesario, la prueba WISC-IV, para complementar la información con las habilidades intelectuales que el niño posee.

Evaluación de los hábitos de estudio

Estudiar es una actividad que se lleva a cabo cuando intentamos adquirir un aprendizaje. Para que los estudiantes de los diferentes niveles de enseñanza –desde el básico hasta el superior–, adquieran los conocimientos que se les proporcionan en la escuela y que les servirán para desenvolverse exitosamente en su vida futura, tanto académica como social y laboral, se requiere por parte de ellos tanto compromiso como motivación para aprender y, desde luego, contar con métodos de estudio eficaces que los capaciten para realizar los trabajos escolares de manera adecuada y responsable.

Para cumplir con tales objetivos es indispensable que los estudiantes sepan cómo estudiar; muchos de ellos aún no toman conciencia de la importancia que tienen los hábitos de estudio en su quehacer educativo y, más aún, consideran que éstos no son indispensables para tener un adecuado desempeño escolar. Por tal motivo, una evaluación de los alumnos al respecto es importante para orientarlos sobre sus fortalezas y debilidades, con el propósito de que establezcan, desarrollen y mejoren sus hábitos de estudios. Existen en el mercado dos pruebas editadas por TEA Ediciones que evalúan este atributo: el Cuestionario de Hábitos y Técnicas de Estudio y el Inventario de Hábitos de Estudio (este último aparecerá publicado este año con normas para México).

El Cuestionario de Hábitos y Técnicas de Estudio, de M. Álvarez González y R. Fernández Valentín, es una prueba de 66 reactivos, orientada tanto a la norma como al criterio, en este último se obtiene el porcentaje de conductas que el evaluado reporta. Esta prueba tiene como finalidad un diagnóstico de aquellos aspectos o factores que inciden en la tarea del estudio; con la información obtenida de la aplicación del cuestionario se puede facilitar el pronóstico de la posible influencia de esos hábitos y técnicas en las tareas de aprendizaje y así diseñar e implementar un programa de

intervención para modificar aquellos aspectos que pueden influir negativamente en el estudio o reforzar aquellos que actúan de forma positiva. En la elaboración de este instrumento se consideraron tres aspectos fundamentales: Condiciones físicas y ambientales, Planificación y estructuración del tiempo y Conocimiento de las técnicas básicas, los cuales son medidos a través de siete escalas:

Actitud general hacia el estudio. Incluye todo lo que hace referencia a la predisposición, interés y motivación hacia el estudio.
Lugar del estudio. Hace referencia a la ubicación física que, de alguna manera, puede contribuir a una mayor concentración y rendimiento en el mismo.
Estado físico del escolar. Alude a las condiciones físicas personales, en cuanto a la situación de su organismo, que le permitan un buen rendimiento en el estudio.
Plan de trabajo. Abarca todo lo que hace referencia a una buena planificación y estructuración del tiempo que se va a dedicar a estudiar al tener en cuenta el número de materias y su dificultad.
Técnicas de estudio. Ofrece pautas de cómo estudiar y recoge los diferentes pasos que deben seguirse para el estudio de un tema o lección.
Exámenes y ejercicios. Se refiere a las pautas que conviene seguir cuando se va a realizar un examen o ejercicio.
Trabajos. Incluye aspectos que se han de tener en cuenta para realizar un trabajo, tales como el esquema inicial, las fuentes de información, el desarrollo y la presentación.

Las puntuaciones obtenidas son el porcentaje de conductas adecuadas que el evaluado reporta realizar en cada una de las escalas, lo cual proporciona elementos para determinar tanto las fortalezas como las debilidades de cada estudiante.

El Inventario de Hábitos de Estudio, de F. F. Pozar, es una prueba diseñada con el propósito de detectar hasta qué punto el estudiante conoce su oficio; el autor conceptualiza al estudio como un profesional y enfatiza dos aspectos: que estudiar constituye un proceso y que las variables de éste son los hábitos de estudio. Esta prueba consta de 90 reactivos y tiene tres objetivos: diagnosticar la naturaleza y grado de los hábitos, actitudes o condiciones con que el estudiante se enfrenta a la tarea de estudio; pronosticar las consecuencias que cabe esperar del influjo de estos hábitos y actuar, a partir del diagnóstico, en la dirección adecuada para modificar los hábitos defectuosos o favorecer la adquisición e incremento de los considerados benéficos.

Este inventario está conformado por cuatro escalas fundamentales y una adicional (Sinceridad) y a su vez, cada una de ellas contiene varias subescalas:

La Escala I, Condiciones ambientales de estudio, evalúa las situaciones tanto físicas como materiales y personales del evaluado, así como su comportamiento en el ámbito escolar, a través de las subescalas: Condiciones ambientales personales, Condiciones ambientales físicas, Comportamiento académico y Rendimiento.
La Escala II, Planificación del estudio, tiene que ver con la organización que el

estudiante hace de su tiempo y sus actividades para dedicarse a estudiar, comprende las subescalas Horarios y Organización del estudio.

La Escala III, Utilización de materiales, estima la forma en que el individuo maneja los recursos que necesita para estudiar, esto es, la búsqueda de información en los libros, la forma y el aprovechamiento de su lectura, la elaboración de resúmenes, el subrayado, las fuentes de consulta, entre otros; las subescalas que comprenden son: Manejo de libros, Lectura y Subrayado-Resúmenes.

La Escala IV, Asimilación de contenidos, evalúa en concreto la forma personalizada en que el estudiante aprende, las técnicas que utiliza para estudiar y aprender, como memorización, resolución de ejercicios, entre otros; sus subescalas son Memorización y Personalización.

Esta prueba proporciona puntuaciones en eneatipos para cada escala; una vez que se obtiene la puntuación directa, ésta se transforma en una de nueve puntos, donde puntuaciones de primero, representan hábitos de estudio adecuados; siete y ocho bueno y nueve muy buenos; puntuaciones de dos y tres corresponden a hábitos de estudio deficientes y una puntuación de 1 a malos hábitos de estudio.

Es recomendable aplicar estas dos pruebas que se acaban de revisar a partir de 6° de primaria, en secundaria, bachillerato y universidad, cuando, una vez descartado que el alumno no tiene deficiencias a nivel intelectual, presente bajo rendimiento escolar; el objetivo sería averiguar si sus calificaciones bajas o reprobatorias se deben a deficientes hábitos de estudio y, sobre todo, detectar dónde están sus fallas o deficiencias para tomar medidas correctivas. Por otra parte, también se recomienda aplicarla como parte de una batería de orientación vocacional, ya que independientemente de la carrera que un individuo vaya a elegir estudiar, es importante que conozca también sus hábitos de estudio para ayudarle a tener éxito en la culminación de sus estudios y en su profesión futura.

Para finalizar con este apartado de pruebas del área académica, revisemos la evaluación para orientación vocacional.

Evaluación en la orientación vocacional

La orientación vocacional constituye un proceso de facilitación, un servicio prestado a un individuo para ayudarlo a elegir y adaptarse a una ocupación, a desarrollar sus aptitudes; en este proceso dinámico, el individuo tiene que descubrir, en primer lugar, cuáles son los aspectos de su persona que son importantes para elegir la ocupación que en el futuro, en su vida adulta, va a desarrollar; en segundo, cuáles son sus aptitudes e intereses; tercero, conocer cuáles son las características deseables para cada una de las áreas ocupacionales y cuarto, cómo compaginar todos estos aspectos, es decir, cómo las características de su personalidad, sus aptitudes, valores y motivaciones se relacionan y complementan con sus intereses vocacionales. Una prueba que resulta útil en este proceso de consejería vocacional es la de Búsqueda Autodirigida (SDS), editada por la editorial El Manual

Moderno.

Búsqueda autodirigida (SDS)

Del inglés *Self-Directed Search* (SDS), es una prueba de orientación vocacional y ocupacional autoaplicable, que también puede ser autocalificable y autointerpretable, puesto que es muy sencilla. Los autores de esta prueba son John L. Holland, Bárbara A. Fritzsche y Amy B. Poweell. Consta de dos Formas, la J, que se diseñó para aplicarse a estudiantes de educación media y los primeros años de educación media superior, y la forma R, que está diseñada para utilizarse con alumnos de nivel medio superior, superior y adultos en general. La muestra mexicana constó, para la forma J, de 1 102 sujetos (de los tres grados de secundaria y adultos) y para la Forma R de 1 583 participantes (de los tres grados de bachillerato, licenciatura y adultos), de 11 diferentes estados de la República Mexicana.

La Búsqueda Autodirigida permite al usuario investigar sus intereses, aspiraciones, actividades, capacidades y posibles ocupaciones en las que pudiera desenvolverse exitosamente en el futuro. Consta de dos cuadernillos: uno es el de Autoevaluación el cual es contestado por el estudiante en 40 min aproximadamente, y el otro es el de Exploración de carreras y ocupaciones, que es donde, a partir de sus respuestas al primero, el alumno puede encontrar carreras y ocupaciones afines a sus características personales.

La SDS parte del supuesto de que existen seis tipos de personalidad ocupacional:

1. Realista (R). Tiene capacidades mecánicas y atléticas, le agrada estar al aire libre y con máquinas, trabajar con objetos más que con personas y se le describe como práctico, natural, persistente, ahorrativo. Algunas profesiones son: mecánico automotriz, topógrafo, geólogo, técnico electrónico.
2. Investigador (I). Tiene capacidades en matemáticas y ciencias, le agrada estar solo y resolver problemas, le gusta trabajar más con ideas que con personas u objetos y se le describe como lógico, curioso, intelectual, callado, independiente. Algunas profesiones son: biólogo, psicólogo experimental o educativo, matemático, físico.
3. Artístico (A). Tiene habilidades artísticas, realiza trabajos creativos originales, es imaginativo, le gusta trabajar más con ideas que con objetos y se le describe como abierto, creativo, independiente, emotivo, impulsivo y original. Algunas profesiones son: actor, caricaturista, coreógrafo, escultor, escritor, fotógrafo.
4. Social (S). Le interesa cómo la gente se relaciona, le gusta ayudar a otros a resolver sus problemas, prefiere trabajar con personas más que con objetos, se le describe como servicial, comprensivo, cálido, cooperativo, amistoso, amable, generoso y paciente. Algunas profesiones son: enfermero, orientador vocacional, terapeuta del lenguaje, trabajador social, psicólogo clínico, psicólogo social, terapeuta ocupacional.
5. Emprendedor (E). Tiene capacidades de liderazgo y para hablar en público, le

interesa el dinero y la política, le gusta influir en la gente, le gusta trabajar más con personas e ideas que con objetos, se le describe como extrovertido, audaz, vigoroso, optimista, sociable, ambicioso. Algunas profesiones son: agente de viajes, gerente de almacén o de empresa, abogado, guía de turistas, político, vendedor.

6. Convencional (C). Tiene capacidades administrativas y matemáticas, le agrada estar en interiores y organizar las cosas, le gusta trabajar con papeles y números, se le describe como práctico, cuidadoso, eficiente, organizado. Algunas profesiones son: secretario, contador, empleado de oficina, cajero de banco.

El evaluado contesta el cuadernillo de Autoevaluación, en el que se le presentan una serie de **actividades**, así como **carreras y ocupaciones** a las que sencillamente tiene que responder si le gustan o no; de **habilidades**, en las que contesta si sabe cómo hacerlas o si le gustaría aprender a hacerlas; además incluye una lista de **capacidades y habilidades** donde el sujeto debe autocalificarse, en una escala del 1 al 7, si a su juicio las posee y en qué medida. Todas estas actividades, habilidades, capacidades, ocupaciones y carreras, están enmarcadas en cada uno de los seis tipos de personalidad ocupacional. Una vez que la persona termina de contestar la prueba, sólo debe contar el número de respuestas donde dijo SÍ en cada tipo ocupacional, así como las puntuaciones otorgadas en las capacidades y habilidades autocalificadas. En la Forma J, el usuario selecciona aquellas dos letras con puntuaciones más altas, en la Forma R, las tres letras; éstas son su Código Holland, es decir, la combinación de los tipos de actividades y capacidades donde el individuo ha marcado que tiene mayor interés y habilidad, lo cual le permitirá, en el cuadernillo de Explorador de carreras y ocupaciones, determinar a qué tipo de actividades laborales corresponde la combinación obtenida. Por ejemplo, una combinación de IE (Investigador-Emprendedor) corresponde a profesiones como psicólogo organizacional, farmacéutico, investigador de servicios públicos. Una combinación IS (Investigador-Social) proporciona profesiones como criminólogo, mercadólogo, psicólogo clínico, abogado penalista. Adicionalmente, para cada una de las profesiones y ocupaciones enlistadas se proporciona el nivel educativo que se necesita para ejercerla, desde un nivel de educación básica hasta grados de especialidad, maestría o doctorado.

Esta prueba se aplica con fines de orientación vocacional cuando el estudiante está por terminar la secundaria, para darle una guía sobre la opción a seleccionar para continuar sus estudios o capacitarse para una profesión; cuando está en el segundo año de bachillerato, para orientarle a elegir el área en la que se va a formar, y cuando va a finalizar el bachillerato, para ayudarle a elegir la carrera a seguir. Se sugiere complementar esta prueba con el Inventario de Hábitos de Estudio, con alguna prueba de personalidad para corroborar que la profesión elegida coincide con sus características de personalidad, y con una investigación adicional sobre el perfil que se requiere para la profesión, el campo de trabajo, el plan de estudios, lugar donde se imparte, entre otros.

Pruebas del área emocional o clínica

En la mayoría de los casos donde se solicita una evaluación psicológica por problemas en el área académica, éstos no siempre se presentan de manera aislada, sino que por lo general están acompañados de otros trastornos de tipo emocional o clínico. Si un niño tiene bajo rendimiento escolar, es difícil no pensar que también presente ansiedad como consecuencia de su bajo aprovechamiento; hay ocasiones en donde un problema de ansiedad puede provocar bajas calificaciones escolares, lo cual se vuelve un círculo vicioso: se tiene ansiedad por los problemas escolares y ésta a su vez origina que aquéllos se agudicen más, lo cual puede generar más ansiedad. Otras veces, el problema clínico es el principal motivo de consulta. Aparte de la ansiedad se pueden presentar otras problemáticas tales como depresión, baja autoestima, problemas de conducta como agresividad, mal comportamiento, inadecuada socialización, etc.; es así como en este apartado revisaremos los tres problemas clínicos que se presentan con mayor frecuencia en niños y adolescentes: ansiedad, problemas de socialización y depresión.

Evaluación de la ansiedad

La ansiedad es un repertorio que suele estar presente con bastante frecuencia entre la población de estudiantes de todos los niveles, desde la enseñanza básica hasta la superior, ya que uno de los estresores en esta etapa de la vida son las exigencias escolares. Cuando los alumnos no tienen los recursos apropiados para encarar dichas demandas, el resultado puede ser la aparición de síntomas psicofisiológicos: apatía, ansiedad, pánico, depresión, agresividad, entre otros.

Así, es importante detectar oportunamente la ansiedad, ya que es una conducta inadaptativa que puede interferir con el desempeño académico y el bienestar emocional. La ansiedad puede ser producida por estresores externos o situacionales, percibidos por el sujeto como peligrosos o amenazantes, como por estímulos internos, tal es el caso de los pensamientos. El estado de ansiedad es una reacción emocional evocada cuando el individuo percibe una situación específica como amenazante, sin importar si en realidad existe algún peligro presente o no.

La ansiedad es un conjunto de respuestas inadaptativas que se presentan en tres dominios de actuación: el cognitivo, el fisiológico y el motor. En el primer nivel se presentan sentimientos de terror, amenaza o catástrofe inminente, inseguridades, así como pensamientos e imágenes altamente específicos de miedo. A nivel fisiológico se producen cambios cardiovasculares, aumento de la actividad electodérmica, del tono músculo esquelético y del ritmo respiratorio, elevación de adrenalina y transpiración, palpitaciones, mareos, dolores de estómago o de cabeza. A nivel motor se distinguen dos tipos de respuestas motoras: directas e indirectas; dentro de las primeras se encuentran inquietud, tics, temblores, gesticulación, tartamudeo, lengua trabada; reducción de la destreza motora, de la habilidad para la discriminación perceptiva, del tiempo de reacción y del aprendizaje y ejecución de tareas complejas; dentro de las segundas, se presentan conductas de evitación y escape de la fuente que provoca la ansiedad.

En México, podemos evaluar la ansiedad en niños y adolescentes por medio de las pruebas CMASR-2 y STAIC, y para adolescentes y adultos con las pruebas IDARE, AMAS e ISRA.

El Cuestionario de Ansiedad Manifiesta en Niños Revisado 2 (CMASR-2), de Cecil R. Reynolds y Bert O. Richmond, es una prueba de 49 reactivos en un formato de respuesta SÍ-NO, diseñada para medir el nivel y la naturaleza de la ansiedad en niños y adolescentes de los seis a los 19 años, posee normas para México y está publicada por la editorial El Manual Moderno. La muestra mexicana para la obtención de normas la conformaron 2 490 niños y adolescentes (1 311 mujeres y 1 179 hombres), de seis a 19 años de edad, de las regiones norte, centro y sur del país, de escuelas tanto públicas como privadas.

Esta prueba proporciona puntuaciones para seis escalas: una de Ansiedad Total, puntuaciones para las escalas: Ansiedad fisiológica, Inquietud y Ansiedad social y dos puntuaciones de validez: Defensividad e Índice de Respuestas Inconsistentes.

Por otra parte, 10 de los 49 reactivos identifican la manifestación de la ansiedad en relación con el desempeño académico, ya que, como mencionábamos antes, todo lo relacionado con el aprendizaje escolar es una fuente común de ansiedad para niños y adolescentes; también existen reactivos para medir otras fuentes de ansiedad de los chicos en esta etapa de su desarrollo como son las relaciones con su familia y con sus compañeros, lo cual puede no sólo afectar su bienestar emocional sino también su desempeño académico.

La Escala de Ansiedad fisiológica está compuesta por 12 reactivos que miden las manifestaciones físicas de la ansiedad; indica la forma en que un niño somatiza la ansiedad. Una puntuación alta indica que éste experimenta las respuestas fisiológicas típicas de la ansiedad: dolores de estómago, de cabeza, insomnio, sudoración.

La Escala de Inquietud contiene 16 reactivos; sugiere que el niño es miedoso, nervioso, hipersensible a las presiones ambientales, está preocupado, internaliza gran parte de la ansiedad y puede sentirse agobiado.

Ansiedad Social consta de 12 reactivos que indagan acerca de la ansiedad en situaciones sociales y de desempeño académico: preocupación acerca de la comparación con otras personas, no ser capaz de vivir de acuerdo con las expectativas de otros individuos significativos en sus vidas, dificultad para concentrarse.

Ansiedad Total es una puntuación global que se compone de los reactivos de las tres escalas anteriores; nos da una referencia para catalogar de manera general al niño como ansioso (puntuaciones T mayores a 60), no ansioso pero sí con predisposición a tener ansiedad o bien manifestando muchas respuestas de ansiedad (puntuaciones T mayores a 55), o no ansioso (puntuaciones T por abajo de 54). Sin embargo, aún cuando la puntuación T no sea significativa, hay que analizar las puntuaciones de las otras escalas para determinar si el niño está manifestando algún tipo de ansiedad.

La Escala de Defensividad nos proporciona información acerca de un informe impreciso o bien que el evaluado ha intentado dar una apariencia “ideal” o demasiado positiva de sí mismo.

En cuanto al Índice de Respuestas Inconsistentes, cuando se dan seis o más pares de respuestas inconsistentes es muy probable que el examinado haya respondido de una manera no veráz a los reactivos de la prueba.

Esta prueba proporciona puntuaciones T y percentiles, utilizándose con mayor frecuencia para la interpretación las puntuaciones T. Si recordamos que éstas tienen una media de 50 y desviación estándar de 10, entonces dentro de la norma se encontrarían las que estuviesen entre 40 y 60; es decir, puntuaciones mayores a 60 representarían que el evaluado está presentando ansiedad.

Se sugiere aplicar esta prueba cuando el niño o adolescente presente algunos síntomas indicativos de ansiedad tales como dolores de estómago, de cabeza, insomnio, pesadillas, nerviosismo, preocupaciones, miedos, entre otros; así como cuando el chico tenga dificultades escolares de cualquier índole, no sólo académicas, sino también sociales.

Existe la versión para adultos de este cuestionario: la Escala de Ansiedad Manifiesta en Adultos (AMAS), para adultos de 19 a 59 años, también publicada por la editorial El Manual Moderno, con tres diferentes versiones: el AMAS-A, para personas de nueve a 19 años, contiene 36 reactivos repartidos en las escalas Ansiedad fisiológica, Inquietud/hipersensibilidad, Preocupaciones sociales/concentración y Mentira; la AMAS-C, para estudiantes universitarios, con 49 reactivos y cinco escalas: Ansiedad fisiológica, Inquietud/hipersensibilidad, Preocupaciones sociales/concentración, Ansiedad ante los exámenes y Mentira; y la AMAS-E, para personas de 60 años en adelante, con 44 reactivos repartidos en cuatro escalas: Ansiedad fisiológica, Inquietud/hipersensibilidad, Temor ante el envejecimiento y Mentira.

El Inventario de Ansiedad Rasgo-Estado (IDARE), de C. D. Spielberger y R. Díaz Guerrero, es una prueba editada por El Manual Moderno adaptada para la población mexicana; su objetivo es medir dos dimensiones de ansiedad: Rasgo y Estado en estudiantes de Bachillerato, universitarios de primer año, universitarios no graduados, adultos normales (pacientes de medicina general y cirugía), así como en aquéllos con problemas neuropsiquiátricos y reclusos.

Este inventario está constituido por dos escalas de autoevaluación separadas que miden dos dimensiones distintas de ansiedad: la Ansiedad Estado y la Ansiedad Rasgo, cada una consta de 20 reactivos en los que el evaluado señala la intensidad de los síntomas que presenta con una escala tipo Likert de cuatro puntos: No en lo absoluto, Un poco, Bastante y Mucho para Ansiedad Estado y Casi nunca, Algunas veces, Frecuentemente y Casi siempre para Ansiedad Rasgo.

La Ansiedad Estado se refiere a una condición o estado emocional transitorio de la persona que se caracteriza por sentimientos de tensión y aprehensión subjetivos percibidos conscientemente, así como por un aumento de la actividad del sistema nervioso autónomo; el evaluado contesta cómo se siente en un momento determinado. Por otro lado, la Ansiedad Rasgo se refiere a las diferencias individuales, relativamente estables, en la propensión a la ansiedad; es decir, a la tendencia de las personas a responder con elevaciones en la intensidad de la ansiedad a situaciones percibidas como amenazantes; el evaluado contesta cómo se siente en general.

Existe una versión de esta prueba para niños y adolescentes de nueve a 15 años de edad, llamada Cuestionario de Autoevaluación Ansiedad Rasgo/Estado en Niños (STAIC, por sus siglas en inglés), de C. D. Spielberger y colaboradores; consta de las mismas escalas, contiene también 20 reactivos por escala y está publicado por TEA Ediciones. Ambas pruebas proporcionan puntuaciones en percentiles para poder determinar si la persona evaluada se encuentra en la norma en cada una de las dos escalas (percentiles 26 a 74), en la norma baja (26 a 35) o por debajo de la norma (menor o igual a 25), lo cual no representaría un problema, o bien en la norma alta (65 a 74) si presenta muchas respuestas de ansiedad, o por arriba de la norma (mayor o igual a 75), lo cual daría indicios para determinar que el evaluado está presentando ansiedad. Se sugiere aplicar esta prueba, al igual que CMAS-R2, cuando se observan algunos síntomas indicativos de ansiedad, como los ya mencionados, así como cuando el chico presenta dificultades escolares.

Otro cuestionario que también mide ansiedad es el Inventario de Situaciones y Respuestas de Ansiedad (ISRA), de J. M. Tobal y A. R. Cano, dirigido a personas a partir de los 16 años con una escolaridad mínima de secundaria y está publicado por TEA Ediciones. El ISRA es un cuestionario de 66 reactivos tipo Likert repartidos en tres escalas: Ansiedad Cognitiva, Ansiedad Fisiológica y Ansiedad Motora; los 22 reactivos de cada escala están planteados o inmersos en diferentes situaciones categorizadas en cuatro diferentes Factores: I) Situaciones que implican evaluación y asunción de responsabilidades, II) Situaciones sexuales y de interacción social, III) Situaciones fóbicas y IV) Situaciones habituales o de la vida cotidiana. Este instrumento no conceptualiza a la ansiedad como un constructo unitario, sino que considera que las respuestas de sistemas distintos en el individuo no siempre varían en la misma forma, por lo que evalúa por separado la frecuencia de respuestas cognitivas, fisiológicas y motoras de ansiedad, ante situaciones susceptibles de generarla; de esta manera, ante cada reactivo, el evaluado tiene que ponderar esa combinación situación-respuesta y anotar con qué frecuencia le ocurre: Casi nunca, Pocas veces, Unas veces sí y otras veces no, Muchas veces o Casi siempre; a mayor frecuencia de ocurrencia mayor puntuación, lo que indica mayor intensidad de la ansiedad medida.

Este instrumento proporciona cuatro medidas de ansiedad: cognitiva, fisiológica, motora y total; esta última es una puntuación global que comprende a las tres anteriores. Después de este análisis de las modalidades de ansiedad que presenta el evaluado, se realiza otro para determinar qué tipo de situaciones son las que desencadenan sus respuestas de ansiedad, y estas medidas se refieren a los cuatro factores enumerados anteriormente.

Esta prueba proporciona dos tablas en percentiles, una para cada una de las modalidades de respuesta y otra para los factores o situaciones; de esta manera, se puede determinar si la persona evaluada está presentando ansiedad y en qué modalidad se manifiesta (puntuación mayor o igual al percentil 75), y ante qué tipo de situaciones se presentan esas respuestas de ansiedad.

Evaluación de problemas de conducta social

Los seres humanos vivimos en sociedad, somos seres sociales y por eso necesitamos adaptarnos a las normas que esa sociedad establece para vivir en armonía y desenvolvernos adecuadamente. La adaptación social es un proceso que consiste en el ajuste del comportamiento del niño o adolescente a sus propios deseos y necesidades pero, al mismo tiempo, éste debe también adaptarse a las circunstancias del entorno en que vive, es decir, a las normas y necesidades de las personas con las que interactúa. El ajuste social correlaciona positivamente con autoestima, amistad y satisfacción de vida y negativamente con victimización verbal, física y relacional. Los niños y adolescentes que se relacionan adecuadamente con sus iguales son bien aceptados en su grupo, tienen una buena autoestima, disponen de mayores recursos de apoyo y mejores sentimientos de bienestar y ajuste psicosocial.

Por otro lado, en la inadaptación social se presenta el desajuste personal, el conflicto o la fricción con las demás personas, el no poder interactuar adecuadamente con ellas, en suma, el comportamiento no social y el comportamiento francamente antisocial.

La conducta social problemática que se presenta puede ser de dos tipos: las conductas antisociales y las conductas no sociales. Dentro de las primeras tenemos aquellas conductas que representan un problemas tanto para los padres como para los maestros, tales como el desacato de las reglas y normas, falta de respeto hacia los demás, agresividad, terquedad y, en casos extremos, las consideradas como pre-delictivas; dichas formas de actuar interfieren tanto con la convivencia armónica como con el rendimiento académico.

Por otro lado, las conductas no sociales son aquellas que si bien no representan un conflicto para los demás, sí lo son para el niño o el adolescente: problemas de aislamiento, retraimiento, timidez, ansiedad social, entre otras; representan un problema ya que, además de una buena comunicación y relación con los padres, una variable que desempeña un importante papel en la adaptación social de niños y adolescentes es el tipo de vínculos que establece con sus amigos; en el ajuste psicológico de un individuo en general es importante contar con el apoyo social de los amigos.

Tres pruebas que permiten ayudar a evaluar los problemas de conducta social que se reportan con mayor frecuencia en niños y adolescentes son la Batería de Socialización para Profesores y Padres (BAS 1-2), la Batería de Socialización (Autoevaluación, BAS 3) y el Cuestionario de Conductas Antisociales y Delictivas; esta última con normas para población mexicana.

La Batería de Socialización (BAS 1-2) es una prueba enfocada a medir socialización en niños y adolescentes, de los seis a los 15 años de edad, en ambientes escolares y extra escolares, desde el punto de vista de sus profesores y padres; los autores son F. Silva Moreno y C. Martorell Pallás, está publicada por TEA Ediciones y tiene dos versiones: BAS 1 para profesores (con 118 reactivos) y BAS 2 para padres (con 114 reactivos); es decir, son los profesores y padres lo que contestan los respectivos cuestionarios para calificar la conducta del niño o adolescente en cuestión, en una escala tipo Likert de

cuatro puntos: Nunca, Alguna vez, Frecuentemente y Siempre.

La prueba proporciona dos perfiles de la socialización: una apreciación global, a través de la Escala Criterial-Socialización, y otro con siete escalas, conformado por cuatro de aspectos positivos o facilitadores de la socialización para medir conductas pro-sociales, y tres escalas de aspectos negativos, perturbadores o inhibidores de la socialización; esto es, que engloba tanto las conductas antisociales como aquellas no sociales. Las escalas positivas son:

Liderazgo, la cual detecta liderazgo en una connotación positiva, adaptativa y aglutinante; popularidad, iniciativa, confianza en sí mismo y espíritu de servicio, en suma, qué tan bien es aceptado el niño por su grupo de compañeros.

Jovialidad, que mide extraversión, sociabilidad, buen estado de ánimo; si el niño tiene relaciones sociales adecuadas y se integra alegremente a su grupo.

Sensibilidad social, que evalúa el grado en que el niño o adolescente es considerado con las demás personas, en especial con los que tienen problemas o son rechazados o apartados.

Respeto-autocontrol, que aprecia el acatamiento de reglas y normas sociales para facilitar la convivencia, el sentido de responsabilidad y en general, la asunción de un rol maduro en las relaciones interpersonales.

Las escalas negativas son:

Agresividad-terquedad, que detecta aspectos tales como conducta impositiva, perturbadora, antisocial, resistencia a las normas, indisciplina, agresión verbal o física e hiperactividad motora.

Apatía-retraimiento, que aprecia retraimiento social, introversión, aislamiento, falta de energía e iniciativa (apatía).

Ansiedad social-timidez, que mide aspectos de la conducta relacionados con la ansiedad social como miedo o nerviosismo y aspectos que están vínculados con la timidez como apocamiento o vergüenza.

Finalmente, la Escala Criterial-Socialización ofrece una visión global del grado de adaptación social que a juicio de sus profesores y padres presenta el evaluado.

Esta prueba nos ofrece puntuaciones percentiles para las ocho escalas evaluadas, en tablas para profesores y padres. Las puntuaciones que representan problema son las que se encuentran por debajo de la norma y en la norma baja en las escalas facilitadoras y en la Criterial-Socialización (percentiles 0 a 25 y 26 a 35, respectivamente) y aquéllas por arriba de la norma y en la norma alta en las escalas negativas (percentiles por arriba de 75 y de 65 a 74, respectivamente). Así, de manera general podemos observar si a juicio de los profesores y padres el niño está bien adaptado socialmente y en caso de que no sea así, qué tipo de problemas tiene: si presenta insuficientes repertorios de conductas pro-sociales, si manifiesta problemas de conducta por aislamiento, timidez o ansiedad

social, o si sus problemas se deben a que tiene conductas antisociales; además de que nos permite comparar la percepción que tienen de su conducta social tanto sus profesores como sus padres y en caso de haber discrepancia, dónde se encuentran y averiguar a qué pueden deberse éstas.

Se recomienda aplicar esta prueba cuando se reporta que el niño no tiene amigos, es tímido, vergonzoso, le da miedo interactuar con los demás, o bien (y es el mayor motivo de consulta), cuando manifiesta problemas de conducta ya sea en casa o en la escuela, como falta de respeto, desobediencia, no acatamiento de las normas, agresividad, terquedad, entre otras.

La Batería de Socialización (Autoevaluación, BAS-3), de los mismos autores y publicada también por TEA Ediciones, es la versión de la prueba anterior pero ahora para evaluar la socialización desde el punto de vista del mismo niño o adolescente (de los 11 a los 19 años); esto es, indaga la percepción que los adolescentes tienen de su propia conducta social. Es una prueba que consta de 76 reactivos que se contestan en un formato de respuesta SÍ-NO. Está compuesta de tres escalas que miden aspectos facilitadores de la socialización; dos miden aspectos negativos, inhibidores o perturbadores y, una evalúa la sinceridad. Las escalas son:

Consideración con los demás, que detecta sensibilidad social y preocupación por los demás, en especial con los que tienen problemas o son rechazados o apartados (correspondería a la de Sensibilidad social en BAS 1-2).

Autocontrol en las relaciones sociales, con dos polos: en el positivo, mide el acatamiento de reglas y normas sociales, y en el negativo, conductas agresivas, impositivas, de terquedad e indisciplina (correspondería a las escalas en BAS 1-2 de Respeto-autocontrol y Agresividad-terquedad).

Retraimiento social, que detecta apartamiento tanto pasivo como activo de los demás, así como aislamiento (corresponde en BAS 1-2 a la escala de Apatía-retraimiento).

Ansiedad social-timidez, que al igual que su correspondiente en BAS 1-2, mide aspectos de la conducta relacionados con la ansiedad social como miedo o nerviosismo y aspectos relacionados con la timidez como apocamiento o vergüenza.

Liderazgo, al igual que su correspondiente en BAS 1-2, detecta ascendencia, popularidad, iniciativa, confianza en sí mismo, espíritu de servicio, aceptación por su grupo de pares.

Al igual que BAS 1-2, esta prueba ofrece puntuaciones percentiles para las escalas evaluadas y, en el mismo tenor, aquellas que representan problema son las que se encuentran por debajo de la norma y en la norma baja en las escalas facilitadoras (percentiles 0 a 25 y 26 a 35, respectivamente) y las ubicadas por arriba de la norma y en la norma alta en las escalas negativas (percentiles por arriba de 75 y de 65 a 74, respectivamente). Una vez obtenidas las puntuaciones, podemos observar la percepción que tiene el chico acerca de sus problemas de conducta social, qué tan aceptado se siente, si se considera que tiene problemas por la falta o la deficiencia en sus habilidades

sociales, o bien, por presentar conductas antisociales. Por otra parte, también es valioso poder contrastar los tres puntos de vista aportados, el del profesor, el de los padres y el del propio chico, para determinar dónde hay coincidencias, discrepancias y averiguar el porqué de estas últimas.

Se recomienda aplicar esta prueba por los mismo motivos ya señalados, siempre que el evaluado tenga 11 años en adelante. Si éste tiene de 11 a 15 años, se deben contrastar sus puntuaciones con las aportadas por sus profesores y padres. Si el adolescente tiene 16 a 19 años, se le puede aplicar esta prueba cuando reporta tener problemas de relación con sus compañeros debido ya sea porque no cuente con las habilidades suficientes para socializar, no tenga amigos, le cueste trabajo entablar amistades, o bien, porque presente conductas que no puede controlar como falta de respeto, agresividad, obstinación, terquedad, entre otras.

El Cuestionario de Conductas Antisociales y Delictivas (A-D) es una prueba que evalúa en adolescentes de los 11 a los 19 años de edad dos aspectos de la conducta social desviada: los comportamientos antisociales y los delictivos; su autor es N. Seisdedos Cubero y está publicada con normas mexicanas por la editorial El Manual Moderno .

La muestra normativa consistió en un grupo de 390 adolescentes de 15 años de edad, elegidos aleatoriamente, adscritos a cuatro escuelas secundarias de la Cd. de Mérida; con fines de comparación se aplicó en esa misma ciudad a jóvenes residentes del Centro de Rehabilitación para Menores de ١٢ a ١٨ años. Este cuestionario consta de 40 reactivos que se reparten por igual en dos dimensiones claramente diferenciadas: la conducta antisocial y aquella que incorpora contenidos o comportamientos que fácilmente caen fuera de la ley (delictivos), en un formato de respuesta SÍ-NO.

Esta prueba proporciona puntuaciones percentiles para cada una de las dos escalas: la de Conductas Antisociales y la de Delictivas; así, es posible determinar si el adolescente evaluado está presentando muchas conductas de alguno de los dos tipos (percentiles de 65 a 74), o bien podría ser catalogado como un adolescente antisocial y, en el peor de los casos, como un adolescente que ya comete actos delictivos (percentiles iguales o mayores a 75).

Es recomendable aplicar esta prueba ya sea cuando los profesores o los padres del adolescente reportan que el chico presenta conductas tales como salir sin permiso, molestar a las personas, romper o tirar cosas ajenas, gastar bromas pesadas, contestar mal a los mayores, negarse a trabajar, pelear, robar, portar algún tipo de arma, tomar alcohol o drogas y/o cualquier otra conducta que pueda ser considerada como antisocial o delictiva. Es importante enfatizar el nivel preventivo que posee esta prueba, para poder tomar las medidas necesarias para evitar que el adolescente que presenta conductas antisociales se involucre en actividades delictivas, y con el que ya lo hace, impedir que siga haciéndolo.

Evaluación de la depresión

La depresión es un síndrome caracterizado por una tristeza profunda, sentimientos de

inutilidad o minusvalía y aislamiento; las características de la depresión pueden agruparse dentro de cuatro dominios psicológicos: afectivo, cognitivo, conductual y fisiológico.

Dentro del primero se encuentran: humor deprimido, sentimientos de tristeza, rechazo, minusvalía, pérdida de la alegría de vivir. El dominio cognitivo se caracteriza por sentimientos de inutilidad, vacío, desesperanza, pesimismo sobre el futuro, pérdida de motivación, las responsabilidades se vuelven tareas enormes y la persona las evita, se presentan también autodenigración y autoacusaciones de incompetencia; en casos extremos pueden presentarse ideas suicidas. En el dominio conductual se observa aislamiento social y baja productividad, limitado nivel de energía, falta de limpieza personal, descuido en el vestir y en el arreglo del pelo, expresión facial de máscara; la persona no es capaz de iniciar actividades nuevas y presenta movimientos lentos y habla reducida y lenta (retardo psicomotor). Dentro del dominio fisiológico puede haber pérdida de apetito y peso (o lo contrario), estreñimiento, perturbaciones del sueño y del ciclo menstrual, pérdida de apetito sexual.

Hasta un 20% de mujeres adultas y 10% de hombres llegan a padecer estos trastornos; son más comunes entre las personas que viven solas y que no tienen a alguien que sea significativo para ellos; también es más probable que se presente en aquellas que tienen familiares cercanos con problemas similares. Las estadísticas internacionales indican un aumento importante de los trastornos depresivos, por lo que se empieza a considerar al siglo XXI como la era de la melancolía; no se trata solamente de la presencia de tristeza, sino que implica desinterés, desaliento y una abrumadora desesperanza vital.

Para el DSM-IV, la depresión es considerada como un trastorno del estado de ánimo en el que la persona se siente deprimida o no es capaz de disfrutar de la vida, presenta problemas de alimentación y sueño, sentimientos de culpa, pérdida de energía, dificultades para concentrarse y pensamientos sobre la muerte. El paciente debe haberse sentido mal la mayor parte del día, casi a diario, durante por lo menos dos semanas; este requisito es necesario para asegurar que pueda diferenciarse de los episodios transitorios bajos que la mayoría de las personas sentimos en ocasiones. Con respecto a esto, Lang y Tisher, autoras de uno de los cuestionarios que vamos a revisar, postulan que la depresión es una respuesta humana normal que varía en intensidad y cualidad; es decir, que todos los seres humanos la hemos experimentado o tenemos la posibilidad de hacerlo en algún momento de la vida.

Los síntomas que se han enumerado anteriormente pertenecen a la depresión en adultos, cuyos síntomas de las reacciones depresivas son diferentes a los niños y adolescentes, en los cuales se pueden presentar problemas de conducta y de delincuencia, fobias escolares, dificultades en el aprendizaje, explosiones asociadas a mal humor, irritabilidad, llantos incomprensibles, tendencia a los accidentes, holgazanería, autodestrucción, alteraciones digestivas y del sueño, temores irracionales, dolores de cabeza de tipo migraña y deficiencias motoras. Se reconoce que la sintomatología de la depresión en la infancia y adolescencia se presenta en seis esferas:

Respuesta afectiva: llantos y sentimientos de tristeza y desgracia.

Autoconcepto negativo: sentimientos de inadecuación, baja autoestima, inutilidad, desamparo, desesperanza, falta de cariño.

Disminución de la productividad mental y de los impulsos: aburrimiento, alejamiento, falta de energía, descontento, poca capacidad para el placer y para aceptar la ayuda, retraso motor.

Problemas psicosomáticos: dolores de cabeza y abdominales, insomnio y otras perturbaciones del sueño.

Preocupaciones: por la muerte, la enfermedad, el yo o los otros, pensamientos suicidas, sentimientos de pérdida (real o imaginaria).

Problemas de agresión: irritabilidad y explosiones de mal humor.

En México podemos evaluar la depresión en niños, adolescentes y adultos con dos pruebas, el Cuestionario de Depresión para niños y el Inventario de Depresión de Beck.

El Cuestionario de Depresión para niños (CDS), de las autoras M. Lang y M. Tisher, es una prueba publicada por TEA Ediciones cuyo objetivo es obtener una evaluación tanto global como específica de la depresión en chicos de ocho a 17 años. La prueba consta de 66 reactivos repartidos en dos grandes escalas, 48 para la Escala Total Depresivo y 18 para la Escala Total Positivo. Los reactivos son leídos por el evaluador al niño o adolescente, el cual debe contestarlos en una escala tipo Likert de cinco puntos: Muy de acuerdo, De acuerdo, No estoy seguro, En desacuerdo y Muy en desacuerdo.

La Escala Total Depresivo consta a su vez de seis subescalas:

Respuesta afectiva. Mide el estado de humor de los sentimientos del sujeto.

Problemas sociales. Se refiere a las dificultades en la interacción social, como aislamiento y soledad.

Autoestima. Sentimientos, conceptos y actitudes del niño en relación con su propia estima y valor.

Preocupación por la muerte o la salud. Alude a los sueños y fantasías del niño relacionados con la muerte y las enfermedades.

Sentimientos de culpabilidad. Se refiere a la autopunición del niño.

Depresivos varios. Cuestiones de tipo depresivo que no pudieron agruparse para formar una identidad.

La Escala Total Positivo consta de dos subescalas:

Ánimo-alegría. Mide la ausencia de alegría, diversión y felicidad en la vida del niño, o su incapacidad para experimentarlos.

Positivos varios. Cuestiones de tipo positivo que no pudieron agruparse para formar una identidad.

Esta prueba proporciona puntuaciones en percentiles tanto para las dos escalas como para cada una de las ocho subescalas. Se obtiene así una puntuación para Total Depresivo, la cual indicaría que el niño presenta depresión (puntuaciones mayores o iguales al percentil 75), o bien, que manifiesta muchas conductas indicativas del trastorno

(puntuaciones en la norma alta, entre los percentiles 65 y 74). A su vez proporciona una puntuación para Total positivo, que indica, en puntuaciones altas, ausencia de elementos positivos en la vida del chico o bien su incapacidad para experimentar placer o diversión; esto es, aunque no sea un niño o adolescente que tenga depresión, sí es un niño o adolescente triste.

Por otra parte, ya sea que el chico tenga depresión, su puntuación esté en la norma alta, o carezca de suficientes elementos positivos en su vida, las puntuaciones de las ocho subescalas nos permitirán determinar aquellas áreas en las que el evaluado tiene problemas; esto es, si tiene depresión, cómo se conforma (p. ej., por baja autoestima, sentimientos de culpabilidad, problemas sociales y preocupaciones por la muerte); si no se le considera deprimido pero sí presenta muchos indicadores de depresión, cuáles son éstos; si bien puede que no esté deprimido, a lo mejor existe un área de su vida que le causa problemas, como no tener suficientes elementos positivos en su entorno, presentar baja autoestima, tener sentimientos de culpa o preocuparse en exceso por la muerte, entre otros.

Las autoras recomiendan la posibilidad de que, además del niño o adolescente, este cuestionario sea contestado por los padres y otras personas significativas, para para calificar las conductas del niño y así obtener una información más amplia acerca de la problemática del chico e incluso aplicarla también a sus hermanos; esto permitiría descartar que la depresión sea una problema que se presente en toda la familia.

Es conveniente aplicar este cuestionario en aquellos casos en que se sospeche depresión: el niño o adolescente parece o es visto como triste, lloroso, aburrido, aislado, apático, con dificultades sociales, con problemas psicosomáticos como dolores de cabeza, de estómago, o dificultades de sueño; si se considera que tiene baja autoestima, sentimientos de inutilidad, sensación de no ser querido; si presenta preocupaciones por la salud o la muerte, o bien manifestaciones anómalas de agresividad. Asimismo, cuando se tenga noticia de otros indicativos como una gran privación, una pérdida de una persona importante para él, un bajo rendimiento escolar, una enfermedad física o una ruptura familiar.

Otro cuestionario que se puede utilizar para evaluar esta problemática es el Inventario de Depresión de Beck (BDI-II), de los autores Aaron T. Beck, Robert A. Steer y Gregory K. Brown, publicado por la Editorial Paidós y cuya finalidad es medir la severidad de la depresión en adultos y adolescentes a partir de los 13 años. Esta versión fue desarrollada para evaluar los síntomas correspondientes a los criterios diagnósticos descritos en el DSM-IV de la APA.

Este inventario es un instrumento de autoinforme compuesto por 21 reactivos en un formato de respuesta tipo Likert de cuatro puntos (de 0 a 3), en el que a cada uno le corresponde una aseveración relacionada con un síntoma de la depresión; el evaluado debe marcar aquella que mejor describa la manera en que se ha sentido durante las últimas dos semanas. Los 21 síntomas y actitudes depresivos elegidos para ser incluidos en el BDI-II se basaron en las descripciones verbales de pacientes y no fueron seleccionados para reflejar ninguna teoría de depresión en particular. Estos síntomas son:

1) estado de ánimo, 2) pesimismo, 3) sensación de fracaso, 4) pérdida de placer (anhedonia), 5) sentimientos de culpa, 6) sentimientos de castigo, 7) disconformidad consigo mismo, 8) acusaciones a sí mismo, 9) ideas suicidas, 10) llanto, 11) agitación, 12) aislamiento social, 13) indecisión, 14) desvalorización, 15) pérdida de energía, 16) cambios en los hábitos de sueño, 17) irritabilidad, 18) cambios en el apetito, 19) dificultades de concentración, 20) cansancio o fatiga y 21) pérdida de interés en el sexo.

El BDI-II se desarrolló como un indicador de la presencia y el grado de los síntomas depresivos que coinciden con los del DSM-IV, y no como un instrumento para otorgar un diagnóstico clínico, por lo que debe ser usado con precaución en caso de ser la única medida de diagnóstico, ya que la depresión puede acompañar a una variedad de trastornos de diagnóstico primario, desde los relacionados con el pánico hasta la esquizofrenia. Adicionalmente, hay que prestar especial atención a las respuestas del reactivo 2 (pesimismo) y del reactivo 9 (pensamiento o ideas suicidas), como indicadores de posible riesgo de suicidio.

El tiempo de aplicación es muy corto, en general, se requieren entre cinco y 10 minutos para contestar el instrumento. Cada reactivo es evaluado según una escala de cuatro puntos, de 0 a tres; la puntuación total se determina mediante la suma de las elecciones para los 21 reactivos, donde la puntuación total mínima es cero y la máxima de 63 puntos. Una vez que se ha calculado la suma de las puntuaciones parciales, se consulta una tabla que nos presenta cuatro categorías diagnósticas de la depresión de acuerdo a la puntuación obtenida, desde Mínima, pasando por Leve y Moderada, hasta Severa.

A pesar de que esta prueba está indicada para evaluar a adolescentes a partir de los 13 años y a adultos, mi recomendación es aplicarla a partir de los 18 años, y de los 13 a los 17 años elegir el CDS o, en todo caso, recurrir a ambas para obtener mayor información. Se sugiere también utilizar este inventario cuando se observa o bien el paciente reporta en la entrevista, alguno de los siguientes síntomas: estado de ánimo deprimido, pérdida del placer, falta de apetito y disminución de peso corporal, dificultades para dormir, fatiga o enlentecimiento, problemas para concentrarse, deseos de morir e ideas suicidas. Los síntomas deben ser lo suficientemente considerables como para causarle al paciente malestar o deterioro en el desempeño laboral o escolar, en la vida social (abandono o discordia) o en alguna otra área de funcionamiento (como las relaciones sexuales). Independientemente de la gravedad o duración de los síntomas, hay que descartar que su estado se debe al uso de sustancias, a enfermedades o si el paciente ha sufrido una gran pena o la pérdida de un ser querido en los últimos dos meses (excepto en este último caso si los síntomas son inusualmente graves).

Si bien se ha proporcionado información sobre las pruebas que hemos venido utilizando con mayor frecuencia de acuerdo a los motivos de evaluación psicológica más solicitados, existen otras que nos pueden proporcionar información útil y, de acuerdo al motivo de consulta, a veces información complementaria. Existen así pruebas de personalidad, de adaptación, y también de aspectos muy específicos como son el afrontamiento, la conducta alimentaria, la asertividad, la satisfacción marital, el Trastorno de Asperger, entre otros.

Dentro de las que evalúan personalidad, tenemos para adultos el Inventario Multifásico de la Personalidad Minnesota-2 (MMPI-2), de Hathaway y Mckinley y, basado en el anterior, el Inventario Multifásico de la Personalidad para Adolescentes (MMPI-A), de Butcher, Graham, Williams y Kaemmer, ambos con normas para México y publicados por la editorial El Manual Moderno, la cual también editó otra prueba de personalidad, el Cuestionario de 16 Factores de la Personalidad (16 PF) de Cattell, Eber y Tatsuoka, dirigido a personas de 16 años en adelante y con normas para México; si bien hoy en día se encuentra discontinuado, próximamente se publicará de nuevo con normas actualizadas. Con base en este instrumento, se desarrollaron el Cuestionario de Personalidad para Niños de seis a ocho años (ESPQ), de Coan y Cattell, el Cuestionario de Personalidad para Niños de ocho a 12 años (CPQ), de Porter y Cattell y el Cuestionario de Personalidad para Adolescentes de 12 a 16 años (HSPQ), de Cattell y Cattell; los tres, publicados por TEA Ediciones.

Para evaluar adaptación, contamos con el Test Autoevaluativo Multifactorial de Adaptación Infantil (TAMAI), de P. Hernández Hernández, publicado por TEA Ediciones, el cual mide Inadaptación Personal, Escolar, Social y Familiar; y con Las Escalas Magallanes de Adaptación, de M. García Pérez y A. Magaz Lago, publicada por la Editorial Albor-COHS de España, las cuales evalúan Adaptación al Padre, a la Madre, a la Escuela, a los Profesores, a los Compañeros y Adaptación Personal.

Dentro de este rubro de adaptación, podemos considerar también dos pruebas elaboradas en México, una es la de Sucesos de Vida, de E. Gómez-Maqueo y C. Durán Patiño, dirigida a adolescentes mexicanos de 13 a 18 años de edad, con la finalidad de detectar problemas emocionales en los jóvenes a través de la evaluación afectiva que ellos mismos hacen de los sucesos de vida estresantes que le han acontecido en las áreas familiar, social, personal, problemas de conducta, logros y fracasos, salud y escolar. La otra es el Inventario Autodescriptivo del Adolescente (IADA), de E. Gómez-Maqueo, B. Barcelata Eguiarte y C. Durán Patiño, dirigido también a adolescentes mexicanos de 13 a 18 años de edad, con el propósito de evaluar las dimensiones familiar, social, escolar, personal y salud. La autodescripción que el joven realiza tiene la finalidad de detectar algunos indicadores de riesgo de conductas problemáticas; ambas pruebas están publicadas por la editorial El Manual Moderno.

Para evaluar afrontamiento podemos utilizar las Escalas de Afrontamiento para Adolescentes (ACS), de E. Frydenberg y R. Lewis, que evalúa tres diferentes estilos de afrontamiento (dirigido a resolver el problema, a disminuir los problemas emocionales que el problema causa y el no afrontamiento), a través de 18 diferentes estrategias (publicada por TEA Ediciones).

Para medir problemas en la conducta alimentaria existen el Cuestionario de Sobreingesta Alimentaria, de O´Donnell y Warren (El Manual Moderno) y el Inventario de Trastornos de la Conducta Alimentaria (EDI-2), de D. M. Garner (TEA Ediciones).

Para evaluar asertividad existe la Escala Multidimensional de Asertividad (EMA) de M. Flores Galaz y R. Díaz Loving, que evalúa asertividad, asertividad indirecta y no asertividad en población mexicana a partir de los 15 años (publicada por la editorial El

Manual Moderno) y el Cuestionario de Aserción en la pareja (ASPA), de M. Carrasco, que evalúa cuatro estilos comunicativos dentro de la pareja: asertivo, sumiso, agresivo-pasivo y agresivo (TEA Ediciones).

Existen también la prueba llamada Perfil del Estrés, de Kenneth M. Nowack, el Inventario de Satisfacción Marital Revisado (MSI-R), de Douglas K. Snyder, y la Escala Gilliam para evaluar Trastorno de Asperger, de James E. Gilliam, los tres publicados por la editorial El Manual Moderno.

Para finalizar, queremos mencionar que éstas son sólo algunas de las diversas pruebas que actualmente existen en México, y que no agotan todos los problemas que se pudieran presentar en la consulta de evaluación psicológica, por lo que es recomendable revisar los catálogos de Instrumentos de Evaluación de las editoriales para saber qué otros instrumentos están publicados que nos pudieran servir para evaluar determinados tipos de problemática y/o complementar la información del caso que estemos tratando.

Índice